The Dissertation Committee for Romana Kristelly
certifies that this is the approved version of the following dissertation:

Structural Basis of RhoA Activation by

Leukemia-associated RhoGEF

Committee:

_____
John JG Tesmer, Supervisor

_____
Marvin L Hackert

_____
Kevin N Dalby

_____
Brent L Iverson

_____
Su Dharmawardhane

Structural Basis of RhoA Activation by Leukemia-associated RhoGEF

by

Romana Kristelly, Mag. Pharm.

Dissertation

Presented to the Faculty of the Graduate School of
the University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of
Doctor of Philosophy

The University of Texas at Austin
August 2004

ACKNOWLEDGEMENTS

I thank my advisor, John Tesmer, for his mentorship, patience, and support. I thank all of my colleagues in the lab, particularly David Lodowski for sharing his encyclopedic knowledge, and Guang Gao for helping with the fluorescence assays. I also thank all of our undergraduate research assistants (especially Brett Earnest) for allowing me to teach them and for teaching me in return. I would also like to thank the members of my dissertation committee for their help in my graduate education and the staff and students of the Department of Chemistry and Biochemistry for providing a welcoming and encouraging atmosphere. I am grateful to my parents, whose foresight, support and involvement have guided me through to this point. Finally, I thank my husband, Marcus, (without his encouragement I would have quit graduate school after my first lecture in Physical Chemistry), and our daughter, Anika, for being such a good girl and letting me finish my work.

Structural Basis of RhoA Activation by

Leukemia-associated RhoGEF

Publication No. _____

Romana Kristelly, Ph.D.

The University of Texas at Austin, 2004

Supervisor: John JG Tesmer

Small GTPases of the Rho family are regulators of cytoskeletal organization, neuronal morphogenesis, and transcription and are implicated in the development of cancer. The activity of Rho GTPases is dependent on their binding to GDP (inactive) or GTP (active) and is tightly regulated by accessory proteins. RhoGEFs (Rho **g**uanine nucleotide **e**xchange **f**actors) activate Rho GTPases by stabilizing their nucleotide-free form via a DH/PH (**D**bl **h**omology/**P**leckstrin **h**omology) domains module. **L**eukemia-**a**ssociated **R**ho**G**EF (LARG) belongs to a subfamily of RhoGEFs, the RH-RhoGEFs, that also contain an RH (**R**egulator of G protein signaling **h**omology) domain N-terminal to

the DH/PH domains and specifically activate RhoA, and not the two other RhoGTPases, Cdc42 and Rac1. RH-RhoGEFs are coupled to G protein-coupled receptor (GPCR) activation because their RH domains interact with $G\alpha_{12/13}$ proteins, which in turn activates their GEF activity. Although the LARG DH domain is sufficient for catalysis the PH domain contributes to nucleotide exchange. To better understand how the LARG PH and RH domains contribute to its activity, and to elucidate the structural determinants of RhoA-specificity, structures of the LARG DH/PH domains alone and in complex with RhoA were determined by x-ray crystallography at 2.1 and 3.2 Å resolution, respectively. To verify the structural findings, mutants were generated and assessed using a fluorescence assay. A novel N-terminal subdomain of the DH domain was discovered, which seems to be important for activity and might be used as a switch to regulate LARG activity. The sequence of this N-terminal extension is conserved throughout the Lbc family of RhoA specific RhoGEFs. PH domain-assisted nucleotide exchange in LARG is dependent on the structural integrity of the junction between the DH and PH domains and not on direct contacts of the PH domain with RhoA. A hydrophobic patch on the PH domain has been discovered, which might be a protein-docking site that is used for regulation of GEF activity by other proteins (e.g $G\alpha_{13}$) or other domains within LARG (e.g. RH domain). Fluorescence assays of fragments including the RH domain showed an inhibitory effect of the RH domain *in vitro*, the structural basis of which will be investigated in future experiments.

TABLE OF CONTENTS

Page

1. Introduction

GTPases and GEFs

GTPases belong to a large family of eukaryotic as well as prokaryotic proteins that have intrinsic GTP hydrolysing ability and share homology with the oncoprotein Ras. GTPases are considered molecular switches because they cycle between an active GTP bound and an inactive GDP bound state. The cycling between the active and inactive state is accompanied by subtle structural changes, which are confined to two regions, called the switch regions.

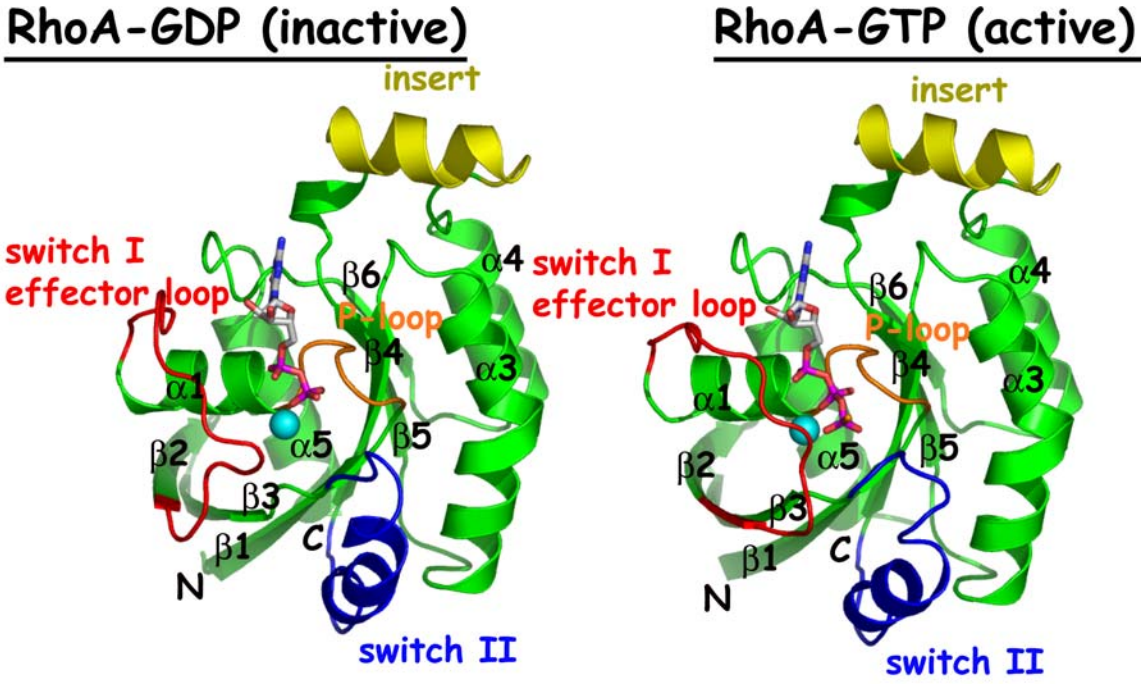

Figure 1.1: Structural differences of inactive GDP-bound and active GTP-bound RhoA are confined to the switch regions. A comparison of GDP and GTP-bound RhoA is shown. The structural changes upon activation are confined to switch I /effector loop (red) and switch II (blue). The P-loop (orange) accommodates the phosphate groups of the nucleotide. A Mg^{2+} (blue sphere) bridges the β- and γ-phosphates. The Rho-specific insertion region (yellow) is implicated in effector binding.

These changes allow activated GTPases to bind to effector proteins. Through this switch mechanism, GTPases control essential cellular processes from membrane transport (Rab, Arf), nuclear transport (Ran), protein elongation (EF-Tu, EF-G), G protein-coupled receptor signaling (Gα subunits), to gene expression, cell cycle progression, and neuronal morphogenesis (Rho). Given the importance of GTPases as regulators of such diverse and essential cellular processes, their activation state is highly regulated. In the resting state of the cell the spontaneous cycling between the inactive GDP bound form and the active GTP bound form proceeds slowly and requires accessory proteins for the cell to be able to respond rapidly to changes in the environment. The transformation of GDP-bound GTPases to the GTP-bound form is catalyzed by guanine nucleotide exchange factors (GEFs) (Figure 1.2). Each family of GTPases is regulated by its respective family of GEFs (Table 1.1). Conversely, the intrinsic hydrolysis of GTP to GDP by GTPases is accelerated by GTPase activating proteins (GAPs).

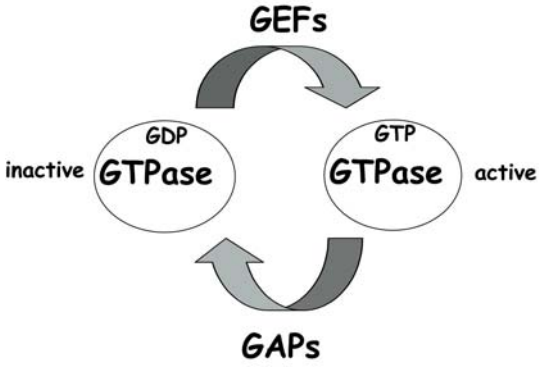

Figure 1.2: GEFs and GAPs control the cycling between inactive GDP-bound and active GTP-bound GTPases. RhoGTPases are molecular switches, which cycle between an inactivated GDP bound state and an activated GTP bound state. The slow intrinsic rates of conversion are accelerated by GTPase activating proteins (GAPs) and Rho guanine nucleotide exchange factors (RhoGEFs).)

GTPase	GEF	GEF structures
Ras	Cdc25 domain proteins	Ras-SOS Cdc25 domain [2]
Rho	Dbl GEFs	**RhoA-LARG** RhoA-Dbs [3] Cdc42-Dbs [4] Cdc42-intersectin [3] Rac-Tiam [5] SOS DH/PH domains [6] Vav DH domain [7, 8] Trio $_{N}$DH domain [9] β-PIX DH domain [8]
	Bacterial RhoGEF SopE	Cdc42-SopE [10]
Ran	RCC1	Ran-RCC1 [11]
Arf	Sec7 domain proteins	Arf-ARNO Sec7 domain [12]
Rab	Mss4	Mss4 [13]
EF-Tu	EF-Ts	EF-Tu-EF-Ts [14]
Gα subunits	GPCR	Rhodopsin [15]

Table 1.1: GTPase families, their respective families of GEFs and known structures of GEFs, either alone or in complex with their effector GTPases.

Although the GEF subfamilies do not share sequence or structural homology, the mechanism by which they promote nucleotide exchange seems to be conserved and for many GEFs it was shown to involve interference with the nucleotide and magnesium binding sites of GTPases and stabilization of the nucleotide free conformation of GTPases. Most kinetic information on the catalysis of nucleotide exchange by GEFs

comes from studies with the Cdc25 domain of Sos [16] which exchanges nucleotides on Ras, EF-Ts [17-19], the GEF for EF-Tu and RCC1 [20], the GEF for Ran, which is not lipid modified. From these studies it was assumed that the nucleotide exchange reaction proceeds through the formation of a quarternary complex of GEF, GTPase, nucleotide, and magnesium, that is followed by expulsion of magnesium and nucleotide and the formation of a high affinity binary complex of GEF with GTPase (Scheme 1). This assumption was supported by several crystal structures of GEFs in complex with their respective GTPases, that all trap the exchange reaction in a nucleotide- and magnesium-free state. In some of these structures (*e.g.* Dbs-RhoA [3], Dbs-Cdc42 [4], Tiam-Rac1 [5], Intersectin-Cdc42 [3]) the only structural basis for nucleotide expulsion seems to be the interference of the GEF with the magnesium binding site, and no interference with phosphates, purine, or ribose were seen. These structures could therefor represent a snap shot of an intermediary state of the exchange reaction and not the high affinity binary complex. Other structures (e.g. SOS-Ras [2], EF-Ts-EF-Tu [14], Sec7-Arf [12]) also exhibit a conformational change in the P-loop that interferes with phosphate binding. These structures confirm the observation from kinetic rate constants that the expulsion of magnesium alone cannot account for the nucleotide exchange rates. It was shown that the rate limiting step in the scheme described above is the release of nucleotide from the quarternary complex [21].

The GEF catalyzed reaction (Scheme 1.1) can proceed in both directions, but the uni-directional exchange of GTP for GDP *in vivo* is assumed to be driven by the cytosolic

excess (3 to 10 fold [22]) of GTP over GDP as well as by proteins that bind to the activated form of the GTPase and shift the equilibrium to the GTP-bound form.

Scheme 1.1: Kinetic scheme for nucleotide exchange reactions, catalyzed by GEFs. GEFs catalyze the exchange of GTP for GDP on GTPases. Although the exchange reaction can proceed in both directions, *in vivo* it is driven by the concentrations of nucleotides and GTPase-binding proteins.

Common structural features of GTPases

GTPases share about 25% sequence identity between, and 40-80% sequence identity within the subfamilies. The minimal highly conserved domain that is responsible for nucleotide binding and hydrolysis is called the G-domain or Ras-like domain, and consists of a six-stranded β-sheet and five α helices. However, GTPases of several subfamilies have insertions or extensions compared to Ras (Figure 1.3). These insertions are necessary for the specific function of these subfamily proteins. For instance, the Gα subunits have an independently folding α-helical domain of about 14 kDa that is thought to function as an internal GAP, by donating and correctly positioning a catalytic arginine [23].

The nucleotide binding pocket is comprised of highly conserved regions that accommodate the phosphate groups (P-loop) and the nucleotide base (^{116}GXKXDL^{121} motif, ^{160}SAK^{162} motif) (Figure 1.4). All GTPases, except for Gα subunits, bind magnesium with high affinity in the GDP and GTP bound state, neutralizing the negative

charges on the β and γ phosphates (Figures 1-4). Removal of the magnesium ion results in nucleotide dissociation and the presence of magnesium inhibits nucleotide dissociation. Gα subunits only bind magnesium with high affinity in the GTP bound state, but not the GDP bound state. It is thought that the βγ subunit prevents nucleotide dissociation.

The structural changes that accompany the cycling between GDP- and GTP-bound form involve two highly conserved regions, the switch I and switch II regions (Figure 1.1). Although the details of the switch mechanism differs in the various GTPase families, the basis for the conformational change is universal and involves contacts of the γ-phosphate oxygens with the conserved residues Thr^{37} and Gly^{62} (Rho nomenclature, Fig. 1.4) in switch I and switch II, respectively. Hydrolysis of GTP causes a release of these contacts and the two switch regions change into a relaxed conformation. The transition from active to inactive GTPase has therefore been termed a "loaded-spring" mechanism [24].

Many GTPases are lipid-modified at their N- or C-termini, which is not only important for membrane localization, but also for protein-protein interactions and regulation [25]. For example, only geranyl-geranylated RhoA can be sequestered in the inactive GDP bound state in the cytosol by interaction with RhoGDI. SOS catalyzes nucleotide exchange more efficiently on prenylated K-Ras [26]. Likewise, RhoGEFs are more active on prenylated Rho than its unprenylated form [27-29]. $G\alpha_{13}$ can only activate Rho-dependent signaling when it is palmitoylated [30].

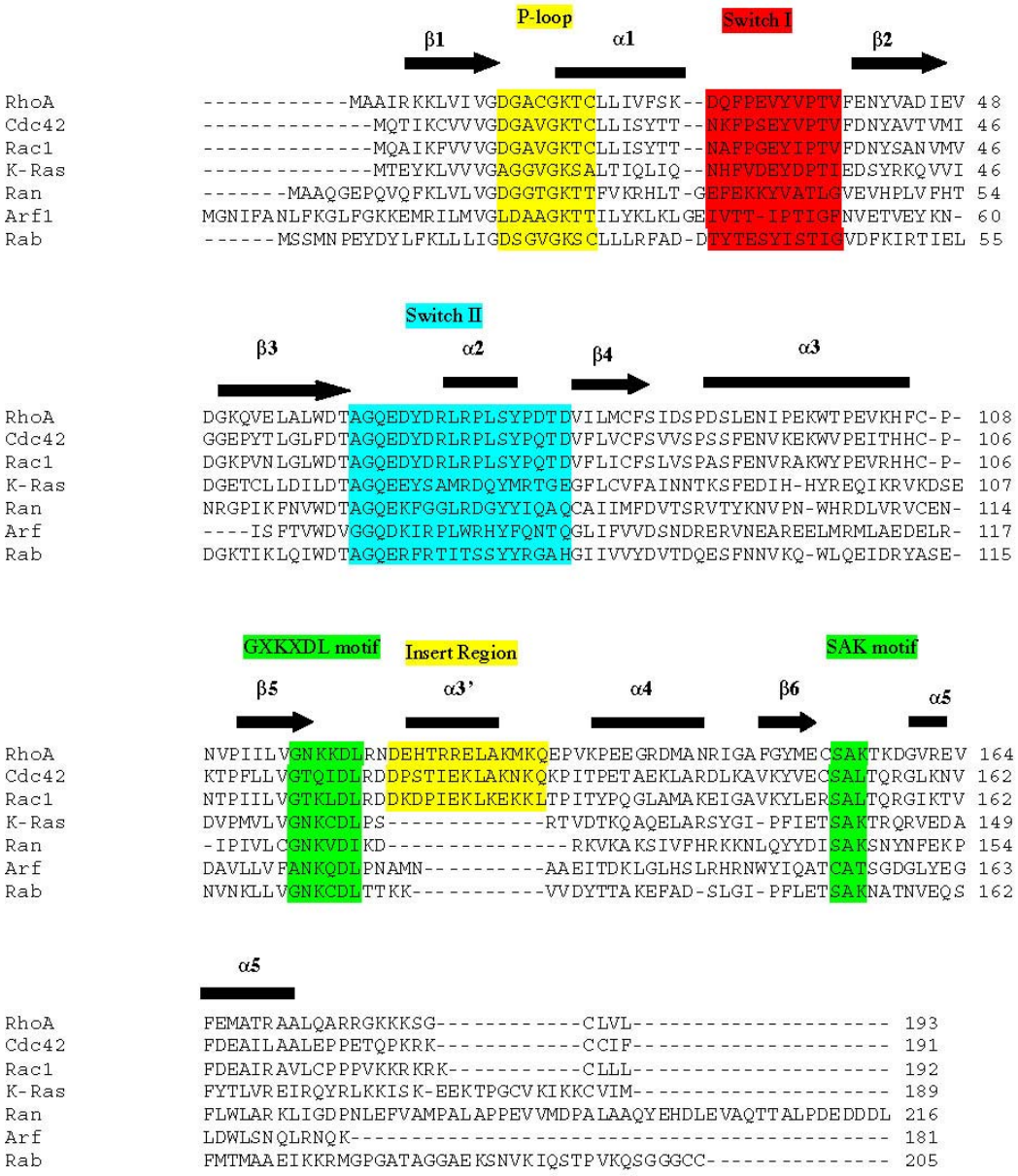

Figure 1.3: Sequence alignment of small GTPases. Highlighd are common structural features of small GTPases and the effector loop of RhoA. The sequence alignment was performed with the program Clustal W [31] and corrected manually according to structural alignment. Accession numbers of sequences used for the alignment are:RhoA3237320, Cdc42 16357472, Rac1 22043610, K-Ras 417590, Ran 32425497, Arfl 4502201, Rab 6679587.

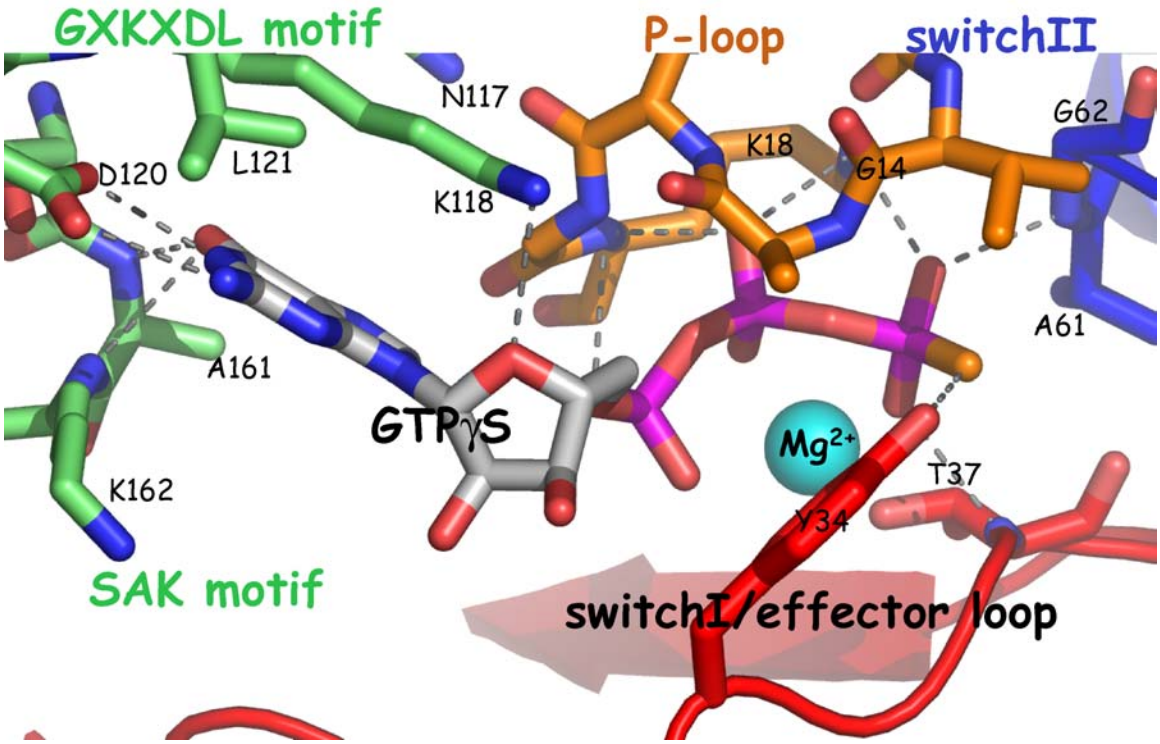

Figure 1.4: The nucleotide binding pocket of RhoA. The nucleotide and Mg $^{2+}$ binding pockets of RhoA are comprised of residues from the switch I and switch II regions, as well as the P-loop, the ^{160}SAK^{162} motif and the ^{116}GXKXDL^{121} motif.

Rho GTPases and RhoGEFs

Rho GTPases are **Ra**s **ho**mology proteins that regulate the actin cytoskeleton, gene expression, and cell cycle progression [32]. Members of the three RhoGTPase subfamilies Rho, Cdc42, and Rac share 40 % sequence identity. The Rho subfamily was the first to be discovered and it has three homologous isoforms, RhoA, RhoB and RhoC, which share 85 % sequence identity. Structurally, Rho GTPases differ from other GTPases by having a 13 amino acid α-helical insertion (Fig. 1.1, 1.3) that is implicated in effector binding [33, 34] and regulation by the regulatory Rho GDIs (**g**uanine nucleotide

dissociation **i**nhibitors), which keep Rho GTPases in the inactive state [35]. Rho GTPases are geranylgeranylated at the N-terminus, which is important for some of their functions [36, 37] and for proper membrane localization.

Rho, Cdc42, and Rac exert different morphological effects on the cytoskeleton [38] and promote transcription through different signaling pathways [39]. However, there is evidence that the signaling cascades of the three individual RhoGTPase subfamilies are functionally linked [40] and the response of Rho GTPases to extracellular stimuli is a result of a complex signaling network involving proteins from all three subfamilies. The Rho GTPase signaling network is also connected to the oncoprotein Ras through SOS, which carries domains that catalyze nucleotide exchange on Ras as well as Rac1 [41] (Table 1). Rho GTPases have been identified as downstream effectors of the oncogene Ras and are ascribed an essential role in Ras-induced tumorigenesis [42].

Rho GTPases are activated by the Dbl (Diffuse B-cell lymphoma) family of RhoGEFs. Dbl was the first RhoGEF of this family to be discovered and was initially identified as an oncoprotein [43]. Since then, approximately 50 related RhoGEFs have been identified in the human genome [44] and together they are the largest family of related oncoproteins. Most RhoGEFs have a tandem Dbl homology/pleckstrin homology (DH/PH) domain module that is responsible for their GEF activity. Individual subfamilies of RhoGEFs contain additional domains, through which they are thought to receive signals from various receptors and regulatory proteins [45]. How and if these additional domains contribute to the exchange activity is not well understood. Recently a bacterial RhoGEF from *Salmonella typhimurium*, SopE, has been discovered, which is

the first known RhoGEF that does not contain a DH domain [46]. SopE has an entirely different structure from the DH/PH domains, although the conformation of Cdc42 bound to SopE is very similar to that of RhoGTPases in complex with DH/PH domains [10].

RH domain-containing RhoGEFs and LARG

Leukemia-associated RhoGEF (LARG) belongs to a family of **r**egulator of G protein signaling **h**omology (RH) domain-containing RhoGEFs (RH-RhoGEFs), which are regulated by $G\alpha_{12/13}$-coupled receptors (Figure 1.5). RH domains were initially identified as negative regulators of G protein signaling processes, because the largest subfamily of RH domain proteins, the RGS proteins, accelerate the intrinsic GTP hydrolysis of Gα subunits by stabilizing its transition state [47]. RGS proteins were therefore termed GTPase activating proteins (GAPs) for the Gα family. Several RH domains with various levels of specificity or promiscuity for the four families of Gα subunits have been discovered.

LARG and the two other members of the RH-RhoGEF family, PDZ-RhoGEF/GTRAP48 and p115RhoGEF/Lfc, selectively activate RhoA and do not act on Rac1 or Cdc42 [48-50]. LARG has an additional amino-terminal PDZ (PSD-95/Dlg/ZO-1) domain (Fig. 1.6) that interacts with the carboxyl-termini of the non-phosphorylated insulin-like growth factor receptor [48] and the transmembrane protein Plexin-B1 [51]. Through Plexin-B1 signaling, LARG is involved in growth cone collapse during nerve cell development. LARG was discovered as an N-terminally truncated form

that was fused to the gene *MLL* (myeloid-lymphoid lineage leukemia) in a patient with acute myeloid leukemia [52] (Fig. 1.6). *MLL* is commonly involved in chromosomal translocations leading to the generation of childhood and drug-induced leukemia [53]. The predicted translation product of the *MLL-LARG* fusion lacks the amino-terminus of LARG including the PDZ domain, and the Zinc-finger domains of *MLL* (Fig. 1.6). It is not known, however, if there is a correlation between aberrant LARG activity and the development of leukemia. Two observations suggest that the intracellular concentration of an N-terminally truncated form of LARG is increased and that this could lead to de-regulation of RhoA activity. Firstly, higher expression levels of amino-terminally truncated forms of LARG compared to full length LARG are seen *in vivo* [54]. Secondly, putative PEST sequences have been identified in the LARG amino terminus [54] and the truncated LARG lacking these sequences could be less prone to degradation by the ubiquitin pathway. Increased intracellular LARG levels might lead to downstream deregulation and this might be a cause for the involvement of LARG in leukemogenic cell development.

Besides its implications in the generation of leukemia, LARG might also play a central role in RhoA-mediated oncogenic transformation by $G\alpha_{12}$ family proteins (*i.e.* $G\alpha_{12}$ and $G\alpha_{13}$). The RH domain of LARG binds to and acts as a GAP on the oncogenic $G\alpha_{12/13}$ subunits. The binding of activated $G\alpha_{13}$ increases the nucleotide exchange activity of the DH/PH domains [55]. Activation of LARG by $G\alpha_{12}$ requires tyrosine phosphorylation of LARG [55]. Investigations of the mechanism by which $G\alpha_{13}$ interacts with RH-RhoGEFs have been undertaken, using p115RhoGEF as a prototype

RH-RhoGEF. Two possible models of activation have been proposed. Based on early experiments that showed an increased GEF activity of a construct lacking the RH domain, an auto-inhibitory role of the RH domain on the DH/PH domains has been assumed [56]. However, later experiments with prenylated RhoA showed decreased activity when the RH domain was removed [57, 58]. Based on studies that identified a binding site for $G\alpha_{13}$ within the C-terminus of LARG, including the DH/PH domains, direct binding of $G\alpha_{13}$ to the DH/PH domains, in addition to the RH domain, might allosterically enhance GEF activity [58].

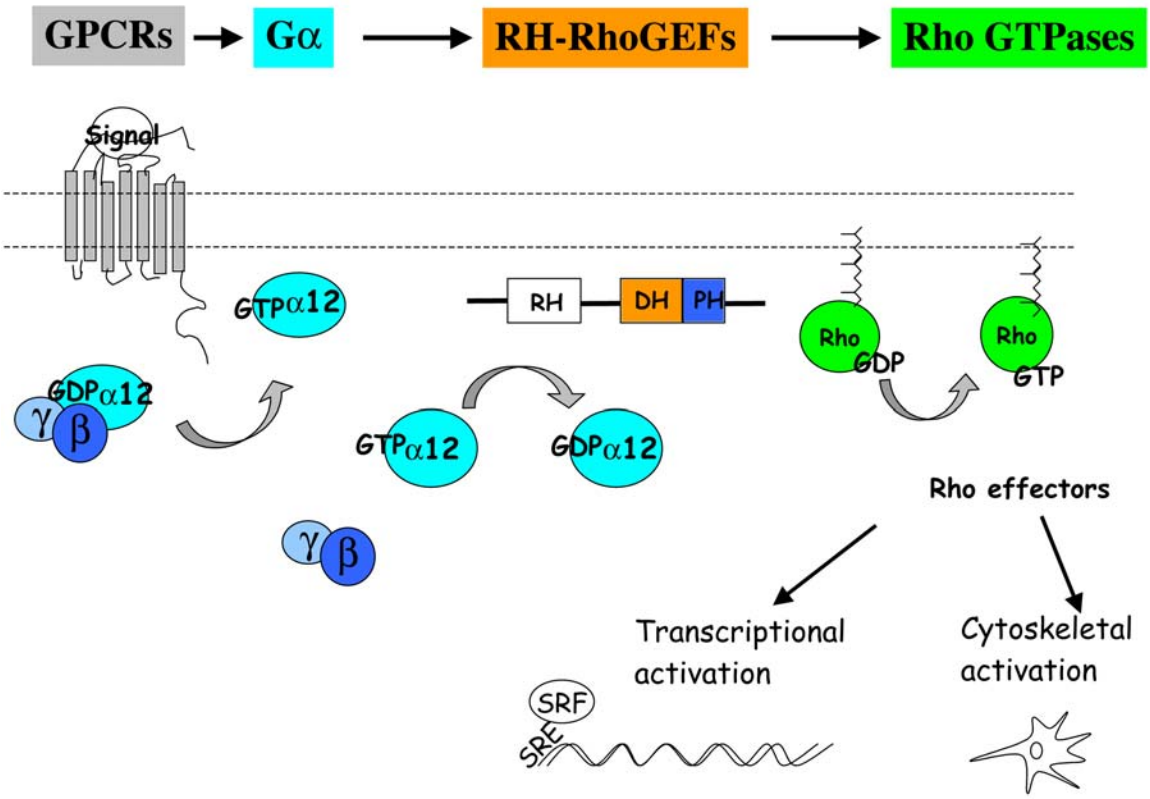

Figure 1.5: Model of signaling from GPCRs to RhoA. A signal (*e.g.* thrombin or lysophosphatidic acid) activates the $G\alpha_{12}$-coupled receptor, which in response activates the $G\alpha_{12}$ family proteins (i.e. $G\alpha_{12}$ and $G\alpha_{13}$). Binding of the activated $G\alpha_{12}$ family proteins to the RH domain of RH-RhoGEFs leads to enhanced activity of the DH/PH domains, which exchange GTP for GDP on RhoA. Activated RhoA, through various effectors, regulates the cytoskeleton as well as transcription from the serum response element (SRE). The molecular mechanism by which the RH, DH and PH domains of RH-RhoGEFs coordinately transduce signals from $G\alpha_{12}$ to RhoA is unknown.

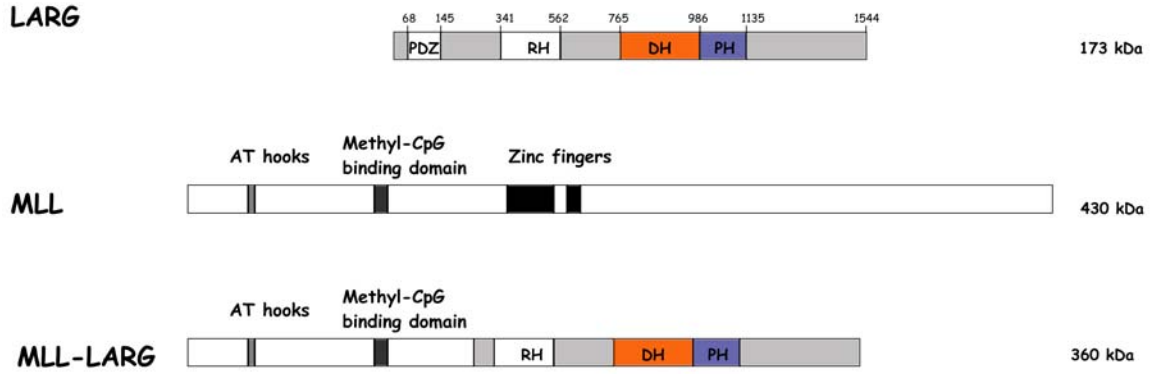

Figure 1.6: LARG has been discovered as a gene fusion with *MLL*. The MLL-LARG fusion protein lacks the PDZ domain of LARG and the Zinc finger domains of MLL.

Goals

The long-term goal of this project is to understand the molecular mechanisms of signaling from oncogenic $G\alpha_{12/13}$ to LARG and RhoA, and in particular how the individual domains of LARG coordinate its activity and regulation. Knowledge gained from this study will not only increase our understanding of the organization of modular proteins but also can ultimately be used for structure based design of drugs that interfere with tumorigenic and metastatic cell development. Furthermore an understanding of the domain structure and organization of LARG will be helpful in explaining a potential role for LARG in the etiology of leukemia.

This research study initially focused on the structural characterization of the interaction between $G\alpha_{12/13}$ and the RH domain of LARG (**Chapter 2**). A structure of either the $G\alpha_{13}$ or $G\alpha_{12}$ subunit by itself would be of great interest because of the oncogenic potential of this Gα subfamily. Chapter 2 describes attempts to express $G\alpha_{12/13}$ in a functional form and with large yields from *E. coli*. Furthermore, attempts are described to purify and crystallize an LARG RH-$G\alpha_{13}$ complex and a complex of the RH with the DH/PH domains. The structure determination of these proteins, however, was hindered by the lack of an appropriate large-scale expression system for $G\alpha_{12/13}$ subunits and by the unsuccessful attempts to crystallize the LARG RH domain. This led to a shift in focus to the interaction of the LARG DH/PH domains with its substrate RhoA and resulted in the structure determination of the DH/PH domains and their complex with RhoA (**Chapter 3**). The focus of Chapter 3 was to answer three questions about the

molecular biology of LARG. (a) What is the structural basis of the specificity for the RhoA subfamily? (b) Can the structures reveal a role for the PH domain in RhoA binding and/or nucleotide exchange? (c) Can the structures reveal possible docking sites for $G\alpha_{13}$ or the RH domain on the DH/PH domains? Structural findings are complemented by steady state kinetic assays of various LARG fragments and DH/PH and RhoA mutants that were designed based on structural analysis (**Chapter 4**). Results from structural and functional analyses taken together reveal novel aspects of the role of the RH and the PH domains in the GEF activity of LARG. Results from Chapters 3 and 4 are discussed and integrated into a model describing the molecular mechanisms involved in signal transduction from LARG to RhoA, and future experiments that need to be performed to validate this model are described (**Chapter 5**).

2. Attempts towards the structure determination of a $G\alpha_{13}$-LARG RH domain complex

2.1. Introduction

The family of $G\alpha_{12}$ subunits has two members, $G\alpha_{12}$ and $G\alpha_{13}$, which share 67% sequence identity. They were first identified in a search for G proteins with transforming potential [59]. Two observations suggested that $G\alpha_{12/13}$ might be upstream regulators of Rho GTPases. First, the expression of constitutively activated mutants of $G\alpha_{12/13}$ and RhoA shows a similar phenotype. Second, the *Clostridium botulinum* C3 exoenzyme, which inactivates Rho GTPases by ADP ribosylation [60], also inhibits $G\alpha_{12/13}$ induced effects [61]. However, the signaling events that transduce signals from $G\alpha_{12/13}$ to Rho GTPases only became evident when it was discovered that $G\alpha_{13}$ directly stimulates nucleotide exchange on the RH-RhoGEF, p115RhoGEF [56]. RH-RhoGEFs contain and RH domain in addition to the DH/PH domain module. Three members of the family of RH-RhoGEFs have been identified: LARG, PDZ-RhoGEF/GTRAP48 and p115RhoGEF/Lsc, which were all found to activate RhoA, but not Rac1 or Cdc42 [49, 50, 54]. The RH domain of p115RhoGEF specifically binds to and acts as a GAP on $G\alpha_{13}$ and to a lesser extent on $G\alpha_{12}$ subunits[62]. Furthermore, $G\alpha_{13}$, but not $G\alpha_{12}$ binding to p115RhoGEF increases its nucleotide exchange activity. Similar result showing differential activity of $G\alpha_{13}$ and $G\alpha_{12}$ have been found for LARG [55]. However, the mechanism by which $G\alpha_{13}$ relays signals to RH-RhoGEFs remains largely unknown. Although early studies with p115RhoGEF indicated that the RH domain auto-

inhibits the DH/PH domains until $G\alpha_{13}$ binds [56], later experiments showed that truncation of the RH domain lead to decreased GEF activity [57, 58]. Furthermore, an additional $G\alpha_{13}$ binding site within a fragment containing the DH/PH domains was discovered, therefore suggesting a complex allosteric mechanism of activation. In this model, the RH domain either directly increases the GEF activity of the DH domain or associates with another region in the RH-RhoGEF, which allosterically increases the DH domain activity [58].

Studies with LARG and PDZ-RhoGEF suggest that separate positive feedback loops might be involved in the overall response of RH-RhoGEFs to extracellular signals. For these proteins it has been shown that additional indirect routes of activation are possible. LARG and PDZ-RhoGEF are both substrates for tyrosine phosphorylation by focal adhesion kinase (FAK) [63], which is a downstream effector of $G\alpha_{12/13}$ and $G\alpha_q$ coupled receptors. In PDZ-RhoGEF it was shown that the region that is responsible for phosphorylation lies C-terminal to the DH/PH domains, and that phosphorylation leads to increased RhoA activation [63]. In a different study, tyrosine phosphorylation of LARG by Tec kinase has been shown to make it responsive to $G\alpha_{12}$ activation, while it does not increase GEF activity in response to $G\alpha_{13}$ [55]. Interestingly, LARG, but not p115RhoGEF or PDZ-RhoGEF, also enhances RhoA activation in response to $G\alpha_q$ [64].

Objectives

This chapter describes work towards the structure determination of the LARG RH domain, $G\alpha_{13}$ and the RH-$G\alpha_{13}$ complex. We focused on the expression of $G\alpha_{13}$ rather

than $G\alpha_{12}$ because functional studies of the interaction of $G\alpha_{13}$ with p115RhoGEF have already yielded some information about the molecular mechanisms involved in activation of RH-RhoGEFs by $G\alpha_{13}$. These structures in conjunction with sequence comparison would help answer the following questions: (1) What are the molecular determinants for specific recognition of $G\alpha_{12/13}$ by the LARG RH domain? (2) What are the structural features of $G\alpha_{13}$ that explain (a) lower GAP activity of the LARG RH domain on $G\alpha_{12}$ compared to $G\alpha_{13}$, and (b) inability of AlF_4^- activated $G\alpha_{12}$ to activate the GEF activity of LARG *in vitro*?

The major obstacle in addressing these questions is the production of suitable quantities of $G\alpha_{13}$, which is necessary for the screening of crystallization conditions. $G\alpha_{12}$ and $G\alpha_{13}$ are mammalian proteins that are post-translationally modified at their N-termini (singly and dually palmitoylated, respectively) and membrane associated. Unmodified $G\alpha_s$, $G\alpha_i$, and a $G\alpha_{it}$ chimera have been expressed in *E. coli* and their structures have been determined [65-67]. Although $G\alpha_{12}$ and $G\alpha_{13}$ have been expressed using the eukaryotic baculovirus expression system in Sf9 insect cells, this purification yields less than 100 μg per liter of cell culture [68]. The development of a large-scale expression system for $G\alpha_{13}$ from *E. coli* would therefore be a major contribution to studies of the molecular and structural biology of these proteins. Several approaches have been taken in this study. (1) $G\alpha_{13}$ was expressed from modified pMALc2X vectors as fusion proteins with maltose binding protein (MBP). MBP has been previously shown to enhance the solubility of its fusion partners and has been termed a molecular

chaperone [64, 69]. (2) An N-terminally truncated form of $G\alpha_{13}$, ΔN47-$G\alpha_{13}$, was expressed. From structural analyses of other Gα proteins this region is known to form a highly mobile α helix that could, in theory, hinder crystallization. (3) Purifications were performed in the presence of detergents to mimic a membrane environment, even though Gα subunits are not post-translationally modified in *E.coli*. (4) Purifications were performed in the presence of AlF_4^- and GDP. In the AlF_4^-•Gα•GDP complex, AlF_4^- forms a square planar complex that is axially coordinated to a β-phosphate oxygen and to a water molecule [66]. A similar coordination (trigonal bi-pyramidal) is assumed for the γ-phosphate at the transition state of the nucleophilic attack by an activated water molecule. AlF_4^- and GDP are therefore used as an analog to mimic the transition state for the intrinsic GTP hydrolysis of Gα subunits. Binding of AlF_4^- or GTP to Gα subunits is associated with movements of the switch regions from a solvent exposed to a more compact, buried environment. (5) Purifications were performed in the presence of the LARG RH domain and AlF_4^-. RH domains are known to bind and stabilize preferentially the AlF_4^- activated form of Gα subunits [70]. (6) MBP-$G\alpha_{13}$ (86 kDa) was co-expressed with the chaperonins GroEL and GroES. GroEL and GroES are the major chaperonins involved in protein folding. The majority of GroEL substrates is smaller than 60 kDa and contains multiple αβ domains. However, successful purification of proteins that are larger than 60 kDa when co-expressed with GroEL/GroES has been reported [71, 72]. Although it is unlikely that the chaperonin is directly involved in proper folding of these proteins, it is possible that accessory proteins that help with folding, are stabilized.

2.2 Methods

2.2.1 Cloning, Expression and Purification of the LARG RH domain

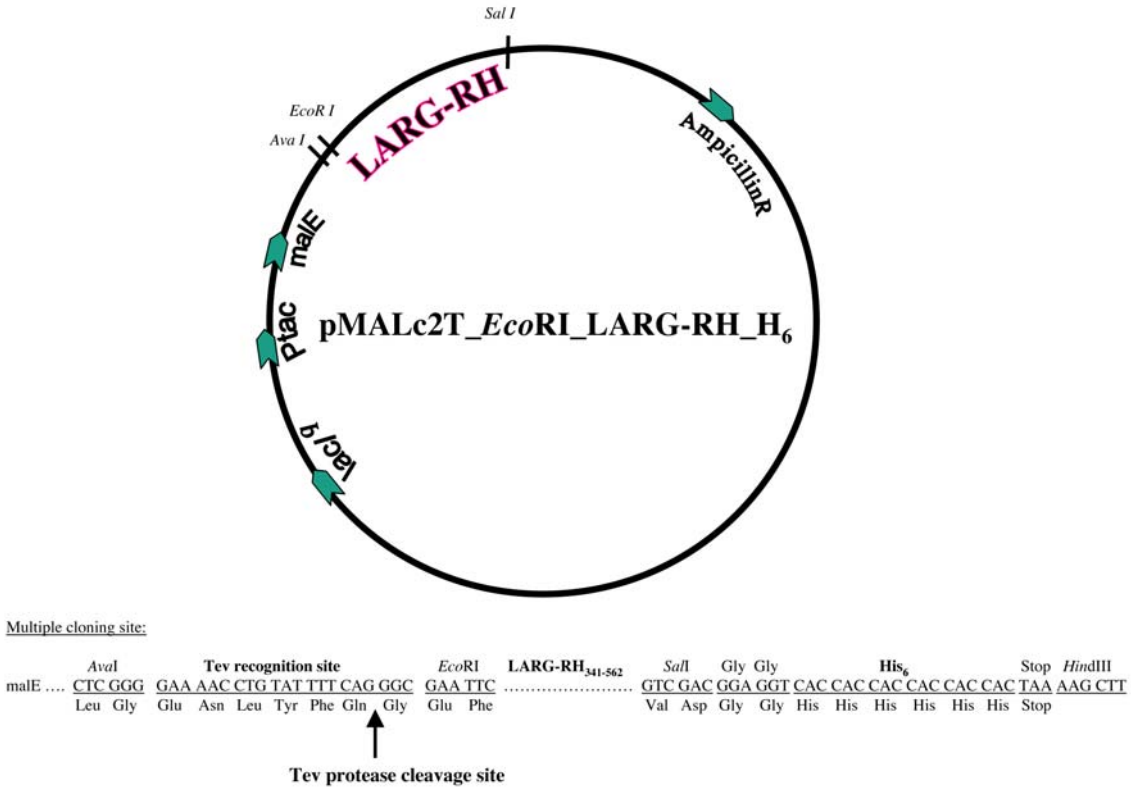

Figure 2.1: Vector map of the expression vector for the LARG RH domain, pMALc2T_EcoRI_RH_H6. The LARG RH domain is expressed as an N-terminal fusion protein with maltose binding protein (MBP) and a C-terminal H_6 tag. After cleavage with TEV protease the RH domain contains additional three amino acids at the N-terminus and four amino acids followed by an H_6 tag at the C-terminus. Arrows indicate the start of the coding regions for the Ptac promotor (Ptac), Lac repressor ($lacI^q$), β-lactamase (AmpicillinR), and maltose binding protein (malE).

The LARG RH domain (residues 341-562) was inserted into a modified pMALc2X vector (New England Biolabs) using *Eco*R I and *Sal* I restriction sites. The modified vector (pMALc2T$_{\text{EcoR I}}$H$_6$) was generated by replacing the factor Xa protease cleavage site with that of tobacco etch virus (TEV) protease, and inserting a sequence that encodes for a C-terminal hexahistidine (H_6) tag followed by a stop codon. The pMALc2X vector was first digested with restriction enzymes *Ava* I and *Eco*R I. 5' phosphorylated oligonucleotides encoding the TEV protease cleavage site (tcggggaaaacctgtattttccagggcg and aattcgccctgaaaatacaggttttcc, Integrated DNA Technologies) were mixed in equimolar amounts, heated to 80 ºC for 10 min, allowed to cool and ligated with the digested pMALc2X vector using T4 DNA ligase (Sigma). To insert the sequence for the C-terminal H_6 tag, the vector was then digested with the restriction enzymes *Sal* I and *Hin*d III, and complementary oligonucleotides (tcgacggaggtcaccaccaccaccaccactaaa and agcttttagaggaggaggaggaggagacctccg, Integrated DNA Technologies) were inserted as described above.

The RH domain fragment of LARG was expressed using the *E. coli* strain Rosetta (DE3) pLysS (Novagen), which carries a plasmid that supplies t-RNAs for six codons rarely used in *E. coli*. Cells were grown in Luria-Bertani (LB) media, supplemented with 0.2% α-D-glucose, 50 μg/ml carbenicillin, and 34 μg/ml chloramphenicol at 37 °C to an OD_{600} of 0.6. Protein expression was induced with 10 μM isopropylthiogalactopyranoside (IPTG) at 20 °C and cells were harvested after 24 hours by centrifugation at 3500 rpm in a Beckman Coulter Avanti J-20 centrifuge using a

JLA 8.100 rotor. The cell pellet was frozen in liquid nitrogen and stored at –80ºC until purification.

Cells were lysed with lysozyme (1 mg/ml) (Sigma) in lysis buffer (20 mM HEPES pH 8.0, 300 mM NaCl, 10 mM β-mercaptoethanol (βME), 10 mM imidazole pH 8.0) supplemented with 0.3 mM EDTA and protease inhibitors (1 μM leupeptin, 1 mM lima bean trypsin inhibitor, 1 mM PMSF, and 1 mM TPCK). Lysates were treated with 20 μg DNaseI (Sigma) per gram of cell pellet and centrifuged for 45 min at 35000 rpm in a Beckman Coulter Optima LE-80K ultracentrifuge, using a Type 45 Ti rotor. The supernatant was loaded onto a drip column containing immobilized Nickel resin (Ni-NTA Superflow, Qiagen), which was pre-equilibrated with lysis buffer. After washing with 5 column volumes of lysis buffer and 5 column volumes of wash buffer (lysis buffer with 20 mM imidazole), protein was eluted in 1 ml fractions with elution buffer (lysis buffer with 250 mM imidazole). Fractions containing the MBP-RH fusion protein were pooled and dialyzed overnight against dialysis buffer (20 mM HEPES pH 8.0, 300 mM NaCl, 5 mM βME) in the presence of 1% (w/w) TEV protease. The RH domain was separated from cleaved MBP by nickel-affinity chromatography as described above for the fusion protein. Fractions containing the RH domain were pooled, concentrated, and loaded onto a size exclusion chromatography column (S200 16/60, Amersham Pharmacia Biotech), which was pre-equilibrated with gel filtration buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 1 mM DTT). The LARG RH domain elutes as a monomer of approximately 25 kDa. Fractions containing the RH domain were pooled, concentrated

to approximately 15 mg/ml and stored at -80 ºC. The typical yield was 7 mg protein per liter of cell culture.

2.2.2 Expression and Purification of $G\alpha_{13}$ from Sf9 cells

$G\alpha_{13}$ was first expressed from Sf9 cells, following a protocol by Kozasa [68]. In this procedure $G\alpha_{13}$ is expressed in a hetero-trimeric complex with $G\beta_1$, and H_6-$G\gamma_2$. The H_6 tag on $G\gamma_2$ allows affinity purification of the complex on a Ni-NTA column. $G\alpha_{13}$ can subsequently be separated from the $G\beta_1H_6$-$G\gamma_2$ heterodimer by activating $G\alpha_{13}$ with the transition state analog AlF_4^-. Dr. Kozasa (University of Illinois at Chicago Medical Center) provided us with the $G\alpha_{13}$, $G\beta_1$, and H_6-$G\gamma_2$ baculoviruses. Sf9 cells were maintained at about $1x10^6$ cells per ml were expanded to 250 ml in IPL-41 medium containing 10% fetal bovine serum in a 500 ml flask. After the cell density reached about 6×10^6 cells per ml (about two to three days) the cell culture was diluted to 1×10^6 cells per ml by expansion to 1 L in four 250 ml cultures in four 500 ml flasks. After two days, cells were further expanded to six 1 L cultures grown in six 2 L flasks in IPL-41 media, substituted with 1% fetal bovine serum and 1% lipid mix (invitrogen). After one day, cells were infected with amplified recombinant baculoviruses encoding $G\alpha_{13}$, $G\beta_1$ and $G\gamma_2$ containing a N-terminal H_6 tag. The virus stock solutions were titered to 4×10^8 pfu/ml and cells were infected with 5 multiplicity of infection units (MOI, plaque forming units (pfu) per cell) for $G\alpha_{13}$, 3.75 MOI for $G\beta_1$ and 2.5 MOI for H_6-$G\gamma_2$. After 48 hours, cells were harvested by centrifugation in a Beckman Coulter Avanti J-20

centrifuge using a JLA 8.100 rotor and cell pellets were frozen in liquid nitrogen. Purification was performed essentially as described by Kozasa [68]. The cell pellet from 4 l was suspended in 600 ml lysis buffer, dounced, and centrifuged at 700 g for 10 min, which removes intact cells and nuclei. The supernatant was then centrifuged at 35000 rpm for 30 min. The supernatant of this spin was removed and saved for further analysis. The pellet, which contains cell membranes and the membrane bound heterotrimeric $G\alpha_{13}\beta_1H_6\text{-}\gamma_2$ complex was re-suspended in 300 ml wash buffer, containing 1 mM $MgCl_2$ and 10 μM GDP. After an additional centrifugation at 35000 rpm for 30 min, the pellet was re-suspended in 200 ml wash buffer. Protein concentration determination by the Bradford method showed 400 mg total protein. Kozasa et al. report a yield of 1.2-2g of protein from an equivalent volume of Sf9 cell culture at this step. Membranes were frozen by pouring them into liquid nitrogen, and stored at –80°C until the next step of purification. Sodium cholate was added to 1% to extract the heterotrimeric $G\alpha_{13}\beta_1H_6\text{-}\gamma_2$ complex from the membranes and the solution was stirred on ice for 1 hr in the presence of protease inhibitors. After centrifuging at 35000 rpm for 40 min, the supernatant was filtered through a 0.2 μm filter and diluted to 1 mg/ml with buffer containing 1 mM $MgCl_2$, 10 μM GDP, and 0.5% Thesit. Subsequently, the solution was loaded onto a NiNTA column and washed with buffer containing an additional 10 mM imidazole. The column was then moved from 4°C to room temperature and washed with buffers containing 0.3 % and 1 % β-octoglucoside, successively. The column was then moved to 30°C and $G\alpha_{13}$ was eluted in 1 ml fractions with buffer containing 1 % β-octoglucoside

and 30 μM AlF_4^-. AlF_4^- acts as a transition state analog and releases $G\alpha_{13}$, while the β_1H_6-γ_2 complex remains bound to the NiNTA resin. The β_1H_6-γ_2 complex was then eluted in 1 ml fractions with buffer containing 150 mM imidazole.

2.2.3 Cloning, Bacterial Expression and Purification of $G\alpha_{13}$

$G\alpha_{13}$ (residues 1 to 377) was amplified from a pCMV5- $G\alpha_{13}$ vector (a gift from Dr. Kozasa, University of Illinois at Chicago, Medical Center) and inserted into the $pMALc2XH_6$ (see 2.2.1) and $pMALc2TH_6$ (see 3.2.1) expression vectors, using *BamH* I and *Sal* I restriction sites. An expression study with the following expression strains BL21 (DE3)RP, BL21 (DE3)RIL, TB1, HMS174 (DE3) pLysS, and BL21 (DE3)RP showed slightly higher expression levels in HMS174, which was then chosen as the expression strain. Later protein expressions were done in the strain Rosetta (DE3) pLysS (Novagen). Protein expression was induced with 10 or 100 μM ITPG at 20°C for 20 hours, and purifications were performed as described for the DH/PH domains in 3.2.1, except that lysis buffer additionally contained 50 μM GDP, 50 μM aluminum chloride, 10 mM sodium fluoride, and either 0.2% Triton X 100, or 0.5% Thesit. Magnesium chloride was added to a final concentration of 5 mM after lysis was completed, as it is known to stabilize bacterial cell membranes. Buffers for affinity and gel filtration columns contained 1% β-octyl-glucoside instead of Triton X100 or Thesit.

Cloning, expression and purification of an N-terminally truncated construct of $G\alpha_{13}$

$G\alpha_{13}$ constructs that lacks 47 amino acids at the N-terminus and do or do not have a stop codon at the C-terminus were cloned into the pMALc2XH_6 plasmid, which was generated as described in 2.2.1, using *Bam*H I and *Sal* I restriction sites. The $G\alpha_{13}$ constructs expressed from these vectors either had an H_6 tag and were purified under various conditions on NiNTA columns, or did not contain an H_6 tag and were purified on amylose columns and used for binding studies with the H_6-tagged RH domain (see next section). Both proteins were expressed from strain HMS174 (DE3) pLysS or Rosetta (DE3) pLysS.

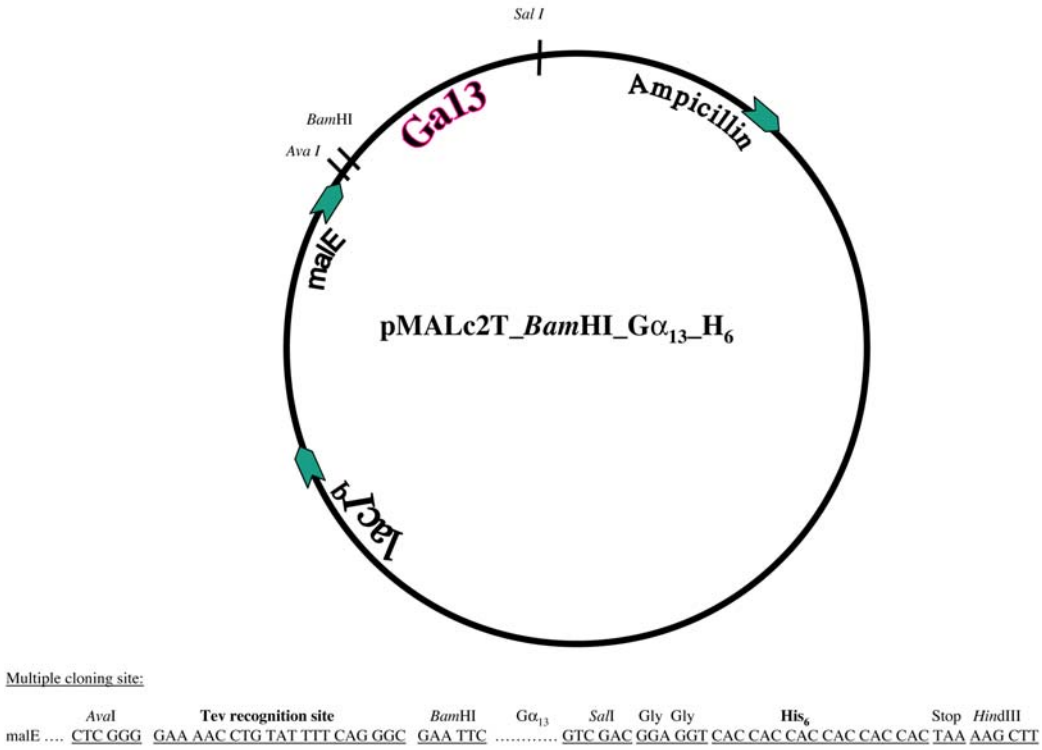

Figure 2.2: Vector map of the expression vector for $G\alpha_{13}$, pMALc2T_BamHI_$G\alpha_{13}$_H_6. $G\alpha_{13}$ is expressed as an N-terminal fusion protein with maltose binding protein (MBP) and a C-terminal H_6 tag. After cleavage with TEV protease $G\alpha_{13}$ contains additional two amino acids at the N-terminus and four amino

acids followed by an H_6 tag at the C-terminus. Arrows indicate the start of the coding regions for the Ptac promotor (Ptac), Lac repressor ($lacI^q$), β-lactamase (AmpicillinR), and maltose binding protein (malE).

2.2.4 Assessment of correctly folded and active $G\alpha_{13}$

MBP-$G\alpha_{13}$-RH complex formation

We attempted to form an MBP-$G\alpha_{13}$•RH domain complex, in order to stabilize $G\alpha_{13}$ during digestion from MBP and to purify this complex for crystallization. A protocol for the formation of a complex between $G\alpha_i$ and RGS4 [73] was followed. MBP-$G\alpha_{13}$ was activated with AlF_4^- by incubation for 15 min at 4°C in complex buffer containing 20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 12 μM GDP, 5 mM $MgCl_2$, 50 μM $AlCl_3$, and 5 mM NaF. Subsequently, 1.5 molar excess of the LARG RH domain was added and the solution was incubated for an additional 15 min on ice. The mixture was subjected to gel filtration in complex buffer or complex buffer containing 1% β-octyl-glucoside on Superdex 75 and 200 columns arranged in tandem and 0.5 ml fractions were collected.

MBP-ΔN47$G\alpha_{13}$-RH complex formation

Complex formation of ΔN47$G\alpha_{13}$ with the LARG RH domain was performed on a NiNTA column (Qiagen). The LARG RH domain, containing a C-terminal H_6 tag was loaded onto a NiNTA column that was pre-equilibrated with buffer containing AlF_4^-. A 1.6 fold molar excess of MBP-ΔN47$G\alpha_{13}$ was incubated with the beads for 12 hours at 4°C. Subsequently, protein was eluted with buffer containing 250 mM imidazole pH 8.0.

Trypsin protection assay

Trypsin protection assays were done in the presence of 30 μM AlF_4^- or 300 μM GTPγS with 0.01-2% (w/w) trypsin (Calbiochem) at 37°C. The time course of digestion could be best observed with 0.01% trypsin. Samples were taken at time points 10, 30 and 60 minutes. Trypsin digests where performed on the MBP-ΔN47$G\alpha_{13}$ fusion protein and on $G\alpha_{13}$ or ΔN47-$G\alpha_{13}$. When trypsin protection assays were performed on the fusion proteins, protein concentrations were high enough for visualization with Coomassie Brilliant Blue staining. However, when performed on the free $G\alpha_{13}$ or ΔN47-$G\alpha_{13}$, western blots had to be performed with a specific $G\alpha_{13}$ antibody (Santa Cruz), which recognizes the N-terminus of $G\alpha_{13,}$ or with a common antibody for Gα subunits (NEN), which recognizes the highly conserved GAGE sequence in the nucleotide binding site.

2.3. Results and Discussion

Expression and Purification of $G\alpha_{13}$ from Sf9 cells

$G\alpha_{13}$ was purified from Sf9 insect cells, following a protocol by Kozasa [68]. Both $G\alpha_{13}$ and $G\beta_1H_6\text{-}\gamma_2$ were barely visible on a gel stained with Coomassie Brilliant Blue (Figure 2.4), and only a total amount of about 130 μg of $G\alpha_{13}$ was detected and 60 μg of $G\beta_1H_6\text{-}\gamma_2$ by the Bradford method. Kozasa reported a yield of 400 μg of $G\alpha_{13}$ from an equivalent size cell culture. Western blots, performed with antibodies that recognize $G\beta_1$ and the N-terminus of $G\alpha_{13}$ appeared to only detect background. A problem with the methodology of the Western blot was likely the reason for the failed detection. It was therefore never convincingly shown that $G\alpha_{13}$ and $G\beta_1H_6\text{-}\gamma_2$ were expressed in Sf9 cells. Possible reasons for low yields of G proteins from Sf9 insect cell membranes are Sf9 cells that were overgrown, or there was incomplete lysis of cell membranes during purification, or deficient virus stock solutions [68]. Furthermore, excessive overexpression of $H_6\text{-}\gamma_2$, which usually shows the highest expression levels in Sf9 cells, can inhibit the expression of Gα subunits [68]. For the screening of crystallization conditions it is estimated that an initial 10 mg of protein are necessary, which would be sufficient for fourty 24 well hanging drop crystal trays, or 960 different crystallization conditions. If a yield of 400 μg per 6 L cell culture could indeed be achieved, expression of $G\alpha_{13}$ from Sf9 insect cells would require one to grow and process at least 150 liters of cell culture, which is not very economical.

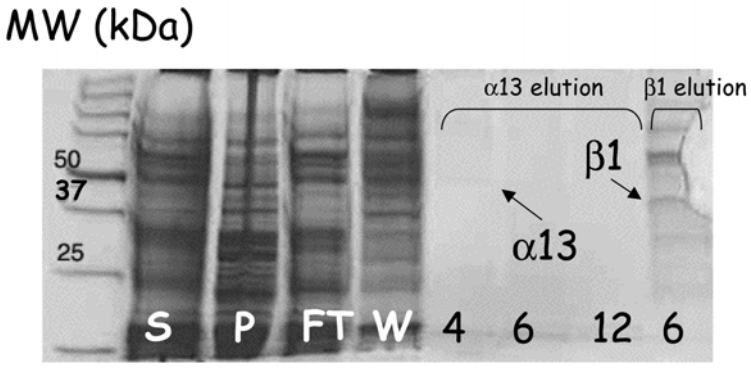

Figure 2.4: Purification of $G\alpha_{13}$ from Sf9 cells. Sf9 cells infected with $G\alpha_{13}$, $G\beta_1$, and H_6-γ_2 baculoviruses were harvested and lysed. The membrane fraction was isolated and the $G\alpha_{13}\beta_1H_6$-γ_2 complex was extracted from the membrane with 1% sodium cholate. After centrifugation the supernatant was loaded onto a NiNTA column. $G\alpha_{13}$ was eluted with buffer containing 30 μM AlF_4^-. The $G\beta_1H_6$-γ_2 complex was eluted with buffer containing 150 mM imidazole. Samples of the eluted fractions were loaded on an SDS-PAGE gel and stained with Coomassie Brilliant Blue. S: supernatant of $G\alpha_{13}\beta_1H_6$-γ_2 cholate extraction from membranes, P: pellet, FT: flow through, W: wash.

Expression and Purification of $G\alpha_{13}$ from *E. coli*

Independent of the various conditions that were explored for expression and purification of $G\alpha_{13}$ from *E. coli*, usable amounts of $G\alpha_{13}$ could not be produced and the following problems were observed: (1) Incomplete and unspecific digestion of fusion proteins with factor Xa or TEV protease (Figures 2.5 and 2.6). This was particularly evident for the MBP-ΔN47-$G\alpha_{13}$ fusion protein, for which complete digestion could not be achieved even with 10% (w/w) TEV protease (Figure 2.6). It is possible that the flexible N-terminal extension of $G\alpha_{13}$ is necessary for proper access of TEV protease to its cleavage site in the MBP-ΔN47-$G\alpha_{13}$ fusion protein, because digestion of full-length

$G\alpha_{13}$ with TEV protease appeared to work better although it was still not complete (Figure 2.5). Typically, MBP fusion proteins are cleavable with 1-3 % TEV protease. Incomplete digestion likely indicates that although some fusion protein might be folded correctly, most of the fusion protein is misfolded and aggregated and access to the protease cleavage site is impeded.

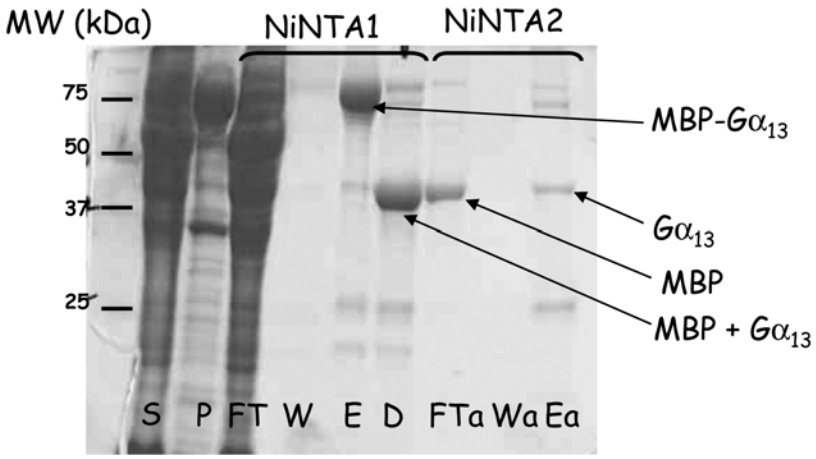

Figure 2.5: Purification of MBP-$G\alpha_{13}$. MBP-$G\alpha_{13}$ was expressed in the *E. coli* strain BL21 (DE3) pLysS carrying a plasmid that encodes for the chaperonins GroEL and GroES. Cells were lysed with lysozyme, centrifuged, and the lysate was applied to a NiNTA column (NiNTA1). MBP-$G\alpha_{13}$ was eluted with 250 mM imidazole, digested with 2 % (w/w) TEV protease, and the digest was subjected to a second NiNTA column (NiNTA2). After washing, protein was eluted from the column with 250 mM imidazole. Samples of the various steps in purification of MBP-$G\alpha_{13}$ were subjected to SDS-PAGE and gels were stained with Coomassie Brilliant Blue. S: supernatant, P: pellet, FT: flow through, W: wash, E: eluate, D: digest.

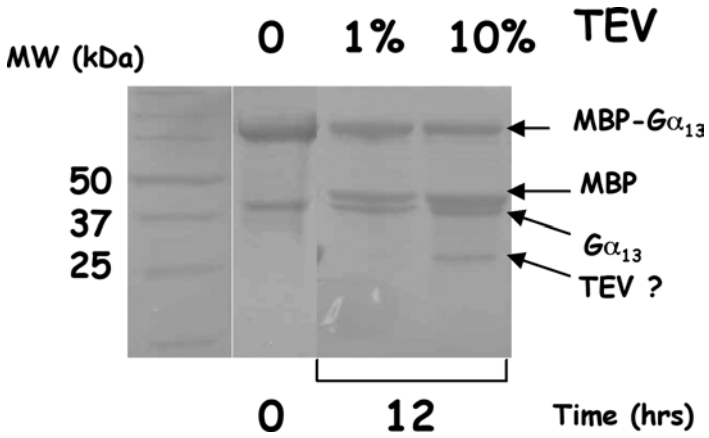

Figure 2.6: Incomplete digestion of MBP-ΔN47-Gα$_{13}$. MBP-ΔN47-Gα$_{13}$ was expressed in the *E. coli* expression strain Rosetta (DE3) pLysS (Novagen) at 20° C with 100 μM IPTG for 19hrs. The cell lysate was subjected to an amylose resin (New England Biolabs) and MBP-ΔN47-Gα$_{13}$ was eluted with buffer containing 10 mM maltose. Digests with the indicated amount of TEV protease for 12 hrs did not result in complete digestion of the fusion protein.

(2) The fusion proteins (86 kDa) eluted from the gel filtration column in the void volume, which is indicative of aggregated protein. (3) Gα$_{13}$ eluted from the gel filtration column at a molecular weight of approximately 80 kDa (Figure 2.7). This could be indicative of a Gα$_{13}$ dimer, which would be around 86 kDa. However, dimerization of Gα subunits has not been observed previously and it is more likely that this represents misfolded protein. (4) Although the yield of the MBP-Gα$_{13}$ fusion protein was high (approximately 5-10 mg per liter cell culture), the yield of Gα$_{13}$ after digestion was low (approximately 100 μg per liter cell culture), which could be equivalent to a small fraction of correctly folded Gα$_{13}$.

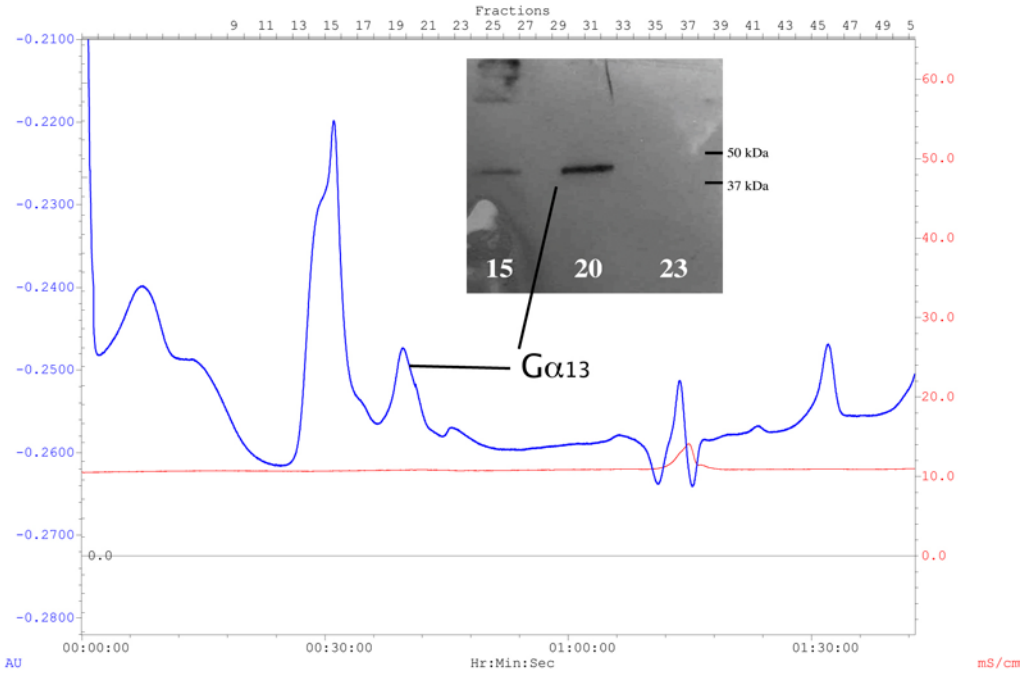

Figure 2.7: Gel filtration of the digested MBP-Gα_{13}-H_6 fusion protein and Western blot of peak fractions. The TEV digest of MBP-Gα_{13}-H_6 was subjected to a NiNTA resin (Qiagen), Gα_{13} was eluted and subjected to two sequentially connected gel filtration columns (Superdex 200, Amersham Pharmacia). Gα_{13} eluted from the column in an 80 kDa peak at a flow rate of 0.5 ml/min. The insert shows a Western blot of fractions from the gel filtration run, probed with an N-terminal antibody (*sc-410*, Santa Cruz).

(5) Two bands just below and just above the 25 kDa standard marker were observed in all purifications of MBP-Gα_{13}, independent of the expression cell line, and the purification conditions (Figure 2.5). Only the band just below 25 kDa was also observed in the purification of MBP-ΔN47Gα_{13} (Figure 2.8). Both bands could not be detected in a western blot with an antibody directed against the N-terminus of Gα_{13}. These proteins bound to NiNTA resin and to a cation exchange resin and could therefore be C-terminal fragments of Gα_{13}. The expressed Gα_{13} construct contains a C-terminal H_6 tag. Gα_{13} has

two domains, a Ras-like domain and an α helical domain, that is inserted between the α1 helix and β2 strand of the Ras-like domain. In addition $G\alpha_{13}$ has an N-terminal helix of 47 amino acids length. A possible digestion product of full length $G\alpha_{13}$ could yield fragments of approximately 27.6 and 14 kDa, corresponding to the Ras-like and α helical domain, respectively. Additional digestion of the N-terminal helix would yield a fragment of 22.6 kDa. The bands just below and just above the 25 kDa marker could therefore be the Ras-like domain of $G\alpha_{13}$ including and excluding the N-terminal extension, respectively. Because these bands are present even during lysis, it would indicate that $G\alpha_{13}$ is partially digested during or even before lysis. (6) Experiments to verify correctly folded and active $G\alpha_{13}$ yielded negative results. Activated $G\alpha_{13}$ did not bind the LARG RH domain. Trypsin protection assays did not show diminished digestion of AlF_4^- or GTPγS activated $\Delta N47G\alpha_{13}$ (Figure 2.9).

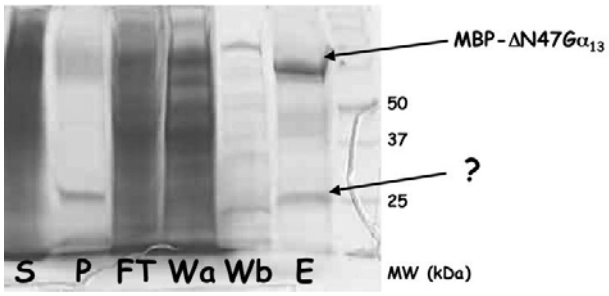

Figure 2.8: Purification of the MBP-ΔN47-$G\alpha_{13}$ fusion protein shows an unidentified band just below 25 kDa. MBP-ΔN47-$G\alpha_{13}$ was expressed in the *E. coli* strain HMS174 (DE3) pLysS. Cells were lysed with lysozyme, centrifuged, and the lysate was applied to a NiNTA column. MBP-ΔN47-$G\alpha_{13}$ was eluted with 250 mM imidazole. Samples of the various steps of the purification were subjected to SDS-PAGE and gels were stained with Coomassie Brilliant Blue. S: supernatant, P: pellet, FT: flow through, Wa and Wb: wash, E: eluate.

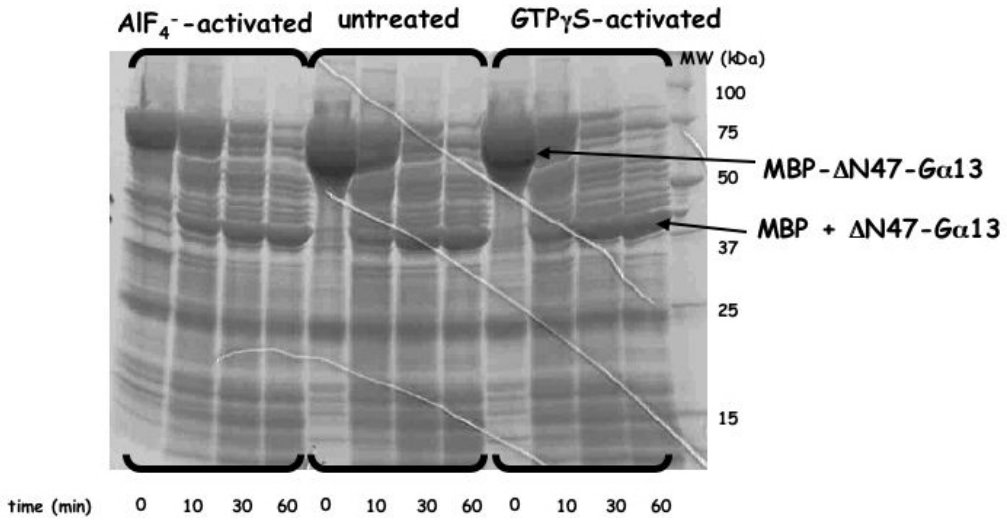

Figure 2.9: Trypsin protection assay of MBP-ΔN47-Gα_{13} fusion protein does not show protection of activated Gα_{13}. MBP-ΔN47-Gα_{13} was incubated with either 30 μM AlF_4^- or 300 μM GTPγS and digests were performed with 0.01 % trypsin (w/w) at 37° C. At the indicated time points samples were subjected to SDS-PAGE and the gel was stained with Coomassie Brilliant Blue. The identical pattern of digested bands under the different conditions indicates that the protein is not protected from trypsin digestion.

Of the various attempts to express and purify functional Gα_{13} from *E. coli* the best purification of full length Gα_{13} could be achieved using 0.5% Thesit for lysis and NiNTA columns, and 10 mM CHAPS for the gel filtration column. 13 mg of MBP-Gα_{13} fusion protein could be retrieved per liter and, ultimately 100 μg of Gα_{13}, which is equivalent to the yield from Sf9 insect cells as reported by Kozasa [68]. Probed with Gα_{13} and MBP antibodies it was shown clearly that the digested band contains both Gα_{13} and MBP, that MBP is filtered out by the NiNTA run, and that the eluate only contained Gα_{13} and MBP-Gα_{13} fusion protein but no MBP (Figure 2.10). Because MBP and Gα_{13} have the same molecular weight and cannot be distinquished on an SDS-PAGE gel, this was useful

information for the interpretation of peaks from gel filtration runs. Strangely, although $G\alpha_{13}$ is appended to an H6 tag, some $G\alpha_{13}$ and MBP-$G\alpha_{13}$ fusion protein could be detected in the flow through of the Ni-NTA column. This could either indicate that the H6 tag was cut off, that the column was overloaded, or that some of the protein was aggregated in a way that prevented binding to the NiNTA resin. $G\alpha_{13}$ eluted from the gel filtration column in the void peak and in a peak at approximately 80 kDa, as probed with $G\alpha_{13}$ antibody (Figure 2.7). Although trypsin protection assays for this protein were performed, no reliable results were obtained. The protein concentrations were too low to be visible on an SDS-PAGE gel stained with Coomassie Brilliant Blue. Therefore, western blots had to be performed, which in this particular experiment gave inconclusive results. However, trypsin digests have been performed on the MBP-ΔN47-$G\alpha_{13}$ fusion protein and did not show a different pattern of digested proteins between AlF_4^-, GTPγS, or GDP-bound $G\alpha_{13}$ (see Figure 2.9). It therefore was concluded that both, the MBP-$G\alpha_{13}$ fusion proteins and $G\alpha_{13}$, retrieved from *E. coli* purifications are not functional, most likely because they are misfolded. This is surprising especially for the MBP fusion proteins because they are expressed in a high yield as soluble proteins. Although disappointing, this result is interesting, because it might indicate the potential of MBP as a chaperone, which can retain even misfolded protein in the soluble fraction during purification. This might be useful for the purification of proteins that can be re-folded to a functional form. In such cases expression of these proteins as MBP fusion proteins could obviate the extraction from inclusion bodies.

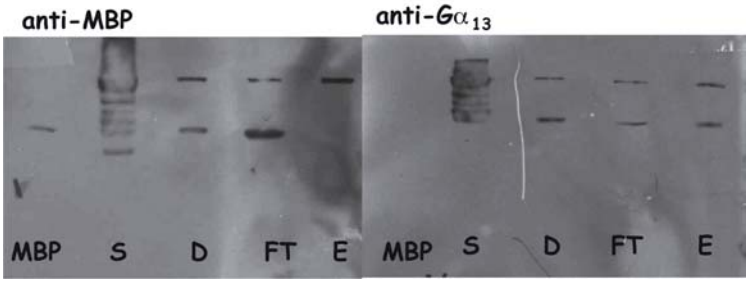

Figure 2.10: Both, MBP and $G\alpha_{13}$ are present in the TEV digest of MBP-$G\alpha_{13}$-H_6 and can be separated by a NiNTA column. Digested MBP-$G\alpha_{13}$-H_6 fusion protein was subjected to a NiNTA resin (Qiagen) and $G\alpha_{13}$-H_6 was eluted with buffer containing 250 mM Imidazol. Samples from various steps of the purification were subjected to SDS-PAGE. Gels were transferred onto a nitrocellulose membrane and probed with anti-MBP (New England Bioloabs) and anti-$G\alpha_{13}$ (*sc-140*, Santa Cruz) antibodies. MBP2 (New England Biolabs) was used as a control for the specificity of the MBP antibody. S: supernatant, D: digest, FT: flow through, E: eluate, MBP: MBP2 protein.

3. The structures of the LARG DH/PH domains and their complex with RhoA

3.1 Introduction

Dbl family proteins

RhoGEFs contain a Dbl homology (DH) domain that is responsible for nucleotide exchange on Rho GTPases and was first identified in the *dbl* oncogene product onco-Dbl, which was isolated from diffuse B-cell lymphoma cells [43]. Almost all RhoGEFs also contain a pleckstrin homology (PH) domain immediately C-terminal to the DH domain. Since the discovery of the oncoprotein Dbl, approximately 50 GEFs of the Dbl family have been identified by virtue of their transforming and metastasis inducing abilities [74]. The Dbl family now represents the largest family of related oncoproteins. The relevance of the Dbl family proteins in other pathologies is evidenced by several examples where RhoGEFs are found at the breakpoints of chromosomal rearrangements (e.g. LARG in MLL [52], FGD1 in Aarskog-Scott syndrome [75], bcr, Philadelphia translocation in CML and AML, [76]). This may indicate an involvement of de-regulated RhoGEFs in the etiology of these diseases.

Common structural features of DH/PH domains

The DH and PH domains are around 250 and 100 amino acids long, respectively. From structural analysis of DH domains (β-PIX [8], Vav [7], Trio [9]), DH/PH domains (SOS [6]), or DH/PH domains complexes with their respective GTPases (Dbs-Cdc42 [4], Dbs-RhoA, ITSN-Cdc42 [3], Tiam-Rac1 [5]) general structural features of the RhoGEF

DH and PH domains were defined. The DH domain is an elongated bundle of six α helical segments (Fig. 3.15). Within segments 1, 2, and 5 there are regions of high sequence conservation that have been termed the conserved regions one, two, and three (CR1, CR2, CR3), respectively. CR1 and CR3 make direct interactions with the GTPase and harbor residues that are responsible for disruption of the magnesium binding site of RhoA. CR2 contributes hydrophobic residues to a core that stabilizes the helical bundle. CR2 might also play a role in oligomerization of RhoGEFs, which is thought to be necessary for RhoGEFs to participate in several interconnected signaling pathways, where presumably large signaling complexes are formed. Evidence for a role of the CR2 in oligomerization comes from studies of Tiam and Dbl. The CR2 of Tiam is involved in the dimer interface in the Tiam-Rac1 crystals [5]. However, the functional significance of dimerization in Tiam is unknown. The CR2 of Dbl is required for its oligomerization, which is essential for Dbl-induced transformation [77].

PH domains are comprised of an anti-parallel flattened β barrel that is capped on one end by a C-terminal α helix (Fig. 3.15, 3.18). Unlike other PH domains, RhoGEF PH domains contain an N-terminal extension that is comprised of an α-helix, which is often an extension of the α6 segment of the DH domain, a β-strand (βN) and a 3_{10} helix (3_{10}N).

Although the common structural features in DH and PH domains have been identified, the details in the various DH/PH domains structures of individual RhoGEF family members have brought more insight into three major unresolved issues in RhoGEF research: (1) the specificity of the various RhoGEFs for their respective

GTPases, (2) the role of the PH domain in nucleotide exchange, and (3) the modes of regulation of RhoGEF activity. A structure of the LARG DH/PH domains will expand knowledge in these areas and lead to a more complete understanding of signaling through RhoGEFs.

Specificity of RhoGEFs

RhoGEFs exhibit a wide range of specificities towards the various RhoGTPase subfamilies. Some RhoGEFs are specific for one GTPase subfamily (*e.g*. LARG is specific for RhoA), others are more promiscuous and exchange on GTPases of two or all three subfamilies (*e.g*. Dbs exchanges on Cdc42 and RhoA, Vav exchanges on Cdc42, RhoA, and Rac1). LARG belongs to the family of RH-RhoGEFs, that are RhoA specific and also include Lbc, Lfc, Net1, XPLN, PDZ-RhoGEF, p115RhoGEF and intersectin, and is called the Lbc RhoGEF family [45]. Intersectin is an outlier in this family because of its specificity for Cdc42 and not RhoA. Understanding the molecular basis for the specificity of the various RhoGEFs will be helpful in designing inhibitors to specifically block signaling cascades of the various Rho GTPase subfamilies.

The specificity determining regions in RhoGTPase and RhoGEFs have been studied by comparing structures of the RhoGEFs intersectin, Dbs, and Tiam with their respective GTPases .[3, 5]. These studies identified variant residues in the β2 and β3 strands of Rho GTPases as the major specificity determining areas that are recognized by specific residues in the α5 helix of the RhoGEF DH. The basis of the specific recognition of Rac1 by Tiam, Cdc42 by intersectin, and the recognition of both Cdc42

and RhoA by Dbs was at least partially explained [3, 5]. However, a DH/PH domain structure of a strictly RhoA-selective RhoGEF has not been available until now.

Role of the PH domain in nucleotide exchange

The conserved positioning of the PH domain immediately C-terminal to the DH domain in nearly all RhoGEFs suggests that it plays an important role. PH domains are traditionally viewed as modules that recruit proteins to membranes and that can bind specific phospholipids [78, 79]. Recent evidence, especially from studies of RhoGEFs [80-83] and GRK2 [84], also points to a function of the PH domain as a protein interaction module for the recruitment to specific intracellular locations and for the docking of regulatory proteins or domains. The Dbl PH domain has also been shown to associate with cytoskeletal proteins [79]. Furthermore, in the non-transforming Dbl protein proto-Dbl, which contains additional 500 amino acids N-terminal of the DH/PH module, compared to onco-Dbl, the N-terminus was shown to specifically bind to the PH domain [85]. This interaction is responsible for keeping Dbl in a non-transforming state. In Vav and SOS, the PH domains exert an inhibitory effect on the DH domain that is relieved upon PIP_3 binding [86, 87]. The role of the PH domain in other RhoGEFs is less well understood. The relative arrangement of PH to DH domains in RhoGEFs has been implicated in dictating the role of the PH domain in the exchange reaction. The most divergent arrangement of the four known structures of DH/PH domains is seen in the SOS DH/PH domains [6]. However, the SOS DH/PH domains are different from the other DH/PH domains because of the inhibitory effect of the SOS PH domain on its DH

domain. In fact, the SOS DH/PH domains are incapable of nucleotide exchange *in vitro*. The arrangement of the state that has been crystallized therefore likely represents the inhibited state of the DH domain.

For three RhoGEFs (Dbs, Trio, and LARG) it has been shown that their DH/PH domains have higher catalytic activities than their DH domains alone [4, 9, 54] and therefore exhibit "PH domain-assisted nucleotide exchange". The mechanism of PH domain-assisted nucleotide exchange has only been explained for Dbs [4, 88]. In Dbs PH domain residues seem to support interactions of the α6 helical segment with the GTPase. However, the structures of RhoGEF PH domains, escpecially their loop regions, are variable and the structural basis for PH domain-assisted nucleotide exchange might vary among the different RhoGEFs.

Regulation of RhoGEFs

Intra-molecular regulation of the DH/PH domains by N- or C-terminal sequences seems to be a common theme in RhoGEFs. Many RhoGEFs (*e.g.* proto-Dbl, proto-Lbc) have been discovered as a result of the isolation of their oncogenic N- or C-terminally truncated forms (*e.g.* onco-Dbl, onco-Lbc). However, several other modes of inter- and intra-molecular regulation of the DH/PH module of RhoGEFs have been identified. Auto-inhibition of the DH domain by the PH domain and alleviation by PIP_3 has been observed in SOS and Vav [86, 87]. Auto-inhibition by N-terminal regions, and release of the auto-inhibition by tyrosine phosphorylation has been reported for Vav [89] and a similar mechanism may also be regulating proto-Dbl activity [85]. Direct interaction of

regulatory proteins with the PH domain has been reported for Dbs, which seems to be regulated by interaction of Rac1 with the PH domain [40]. G protein βγ binding to the PH domain is also implicated in RhoGEF regulation [90]. The formation of homo- or hetero-oligomers has also been proposed as a means of RhoGEF regulation, especially when they participate in large signaling complexes. There are examples of regulation of GEFs by oligomerization. Dbl forms oligomers through the CR2 in the DH domain, which, although not necessary for its *in vitro* exchange activity is necessary for its full catalytic activity *in vivo* [77]. The Tiam-Rac1 complex crystallized as a dimer (worthylake). The dimer interface includes an extensive hydrophobic surface area in the CR2, which also points to the possibility that tiam is regulated by oligomerization.

Although there are no reports on regulation of the RH-RhoGEF PH domain by phospholipids or inhibition of the DH domain by the PH domain, auto-inhibition by N-terminal sequences has been implicated in the regulation of p115RhoGEF. It has been shown that deletion of the N-terminal region of p115RhoGEF, including the RH domain enhanced GEF activity on unprenylated RhoA [56]. For LARG it has been shown that if the RH domain is overexpressed signaling to the serum response element (SRE) is diminished [91]. This might indicate that the RH domain associates with and regulates the DH/PH domains. However, it is unknown which regions in the DH/PH domains associate with the RH domain.

Direct association of the DH/PH domains with $G\alpha_{13}$ has also been proposed as a mode of regulation for the RH-RhoGEFs. Studies with p115RhoGEF identified a C-terminal region including the DH/PH domains as an association site for $G\alpha_{13}$ [58].

Structural analysis of the LARG DH/PH domains will aid in identifying intra-molecular interaction sites for the RH domain and inter-molecular interaction sites for $G\alpha_{13}$ and help to establish a model for regulation of the DH/PH domains of RH-RhoGEF family proteins.

Objectives

The objectives of the present study were to determine the structure of the LARG DH/PH domains and their complex with RhoA and to answer three questions about the molecular biology of LARG: (1) What is the structural basis of the specificity for the RhoA subfamily? (2) What is the structural basis for PH domain-assisted nucleotide exchange in LARG? (3) Can the structures reveal possible modes of regulation of the DH/PH domains, in particular possible docking sites for $G\alpha_{13}$ or the RH domain?

Described in the following section are the cloning, expression and purification of the LARG DH/PH domains and its substrate RhoA and crystallization and structure determination of the DH/PH domains and the DH/PH-RhoA complex.

3.2 Methods

3.2.1 Cloning, Expression and Purification of the LARG DH/PH domains

A fragment of human LARG encoding the DH/PH domains (residues 765 to 1138) was amplified from the pcDNAmycI-Full382 vector (a gift from Dr. T. Kozasa). This vector encodes the LARG gene *KIAA0382* (accession number 2224704), which has a single nucleotide polymorphism (SNP) in the DH/PH sequence compared to the other LARG gene in the NCBI data base, *ARHGEF12* (accession number 34395525). This SNP results in the translation of residue 973 as phenylalanine instead of tyrosine. The amplified DH/PH fragment was inserted into a modified pMALc2X vector (New England Biolabs) using *Bam*H I and *Sal* I restriction sites (Figure 3.1.) [92]. This modified vector (pMALc2T$_{Bam\text{HI}}$_H$_6$) was generated by replacing the factor Xa protease cleavage site with that of tobacco etch virus (TEV) protease, and inserting a sequence that encodes for a C-terminal hexahistidine (H_6) tag followed by a stop codon. The pMALc2X vector was first digested with restriction enzymes *Ava* I and *Bam*H I. 5' phosphorylated oligonucleotides encoding the TEV protease cleavage site (tcggggaaaacctgtattttccagg and gatccctgaaaatacaggttttcc, Integrated DNA Technologies) were mixed in equimolar amounts, heated to 80 ºC for 10 min, allowed to cool and ligated with the digested pMALc2X vector using T4 DNA ligase (Sigma). To insert the sequence for the C-terminal H_6 tag, the vector was then digested with the restriction enzymes *Sal* I and *Hin*d III, and complementary oligonucleotides (tcgacggaggtcaccaccaccaccaccactaaa and agcttttagaggaggaggaggaggagacctccg,, Integrated DNA Technologies) were inserted as described above. In the process of cloning the LARG DH/PH domains a silent mutation

was introduced at residue Asp[765]. The LARG DH/PH fragment, expressed from the pMALc2T$_{BamHI}$_LARG-DHPH_H_6 vector is a fusion protein with maltose binding protein (MBP) at the N-terminus and a H_6 tag at the C-terminus. After treatment with TEV protease, the H_6-tagged DH/PH domains have a molecular weight of approximately 43 kDa.

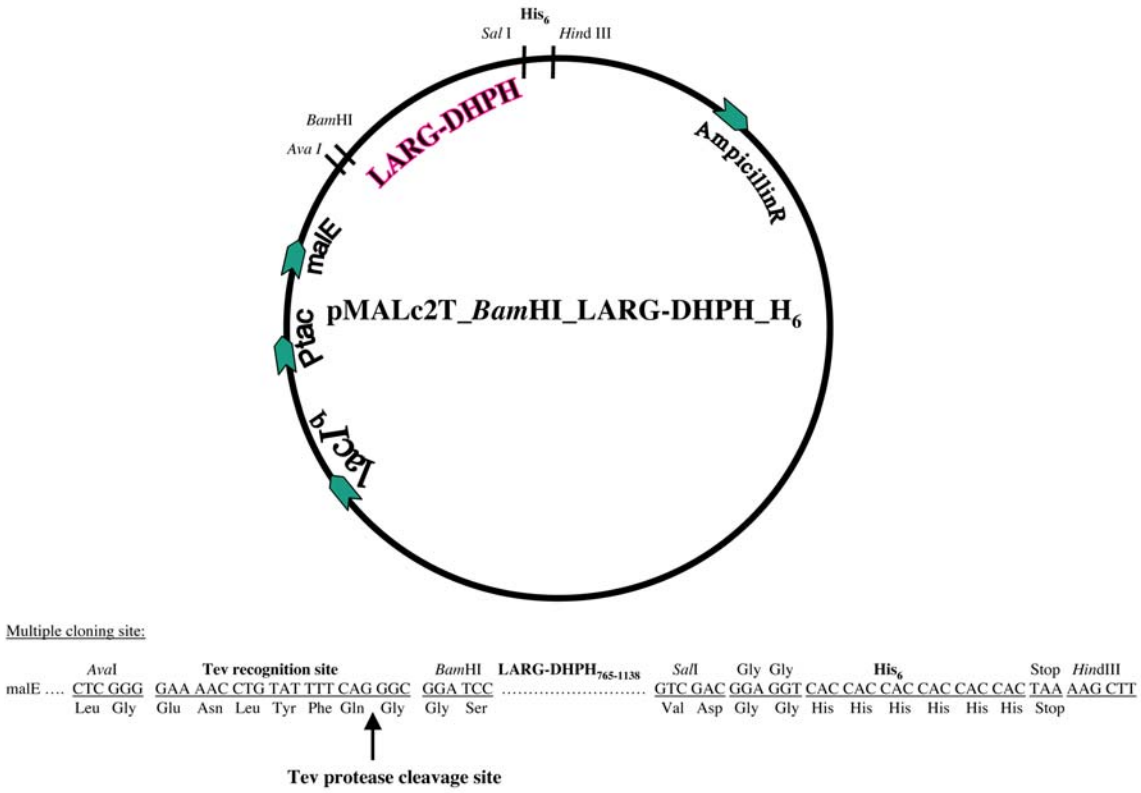

Figure 3.1: Vector map of the expression vector for the LARG DH/PH domains, pMALc2T$_{BamHI}$_LARG-DHPH_H_6. The LARG DH/PH domains were expressed as N-terminal fusion protein with maltose binding protein (MBP) and a C-terminal H_6 tag. After cleavage with TEV protease the DH/PH domains contain additional two amino acids at the N-terminus and four amino acids followed by an H_6 tag at the C-terminus. Arrows indicate the start of the coding regions for the Ptac promotor (Ptac), lac repressor ($lacI^q$), β-lactamase (AmpicillinR), and maltose binding protein (malE).

The DH/PH fragment of LARG was expressed and purified as described for the LARG RH domain in 2.2.1 and in [92], except that the MBP-DH/PH fusion protein was digested with 1.5% (w/w) TEV protease (Fig 3.2). The LARG DH/PH domains eluted from the gel filtration column with an apparent MW of 33 kDa (Fig. 3.3), and subsequently were concentrated to approximately 10 mg/ml and stored at -80 ºC. The typical yield was 15 mg protein per liter of cell culture.

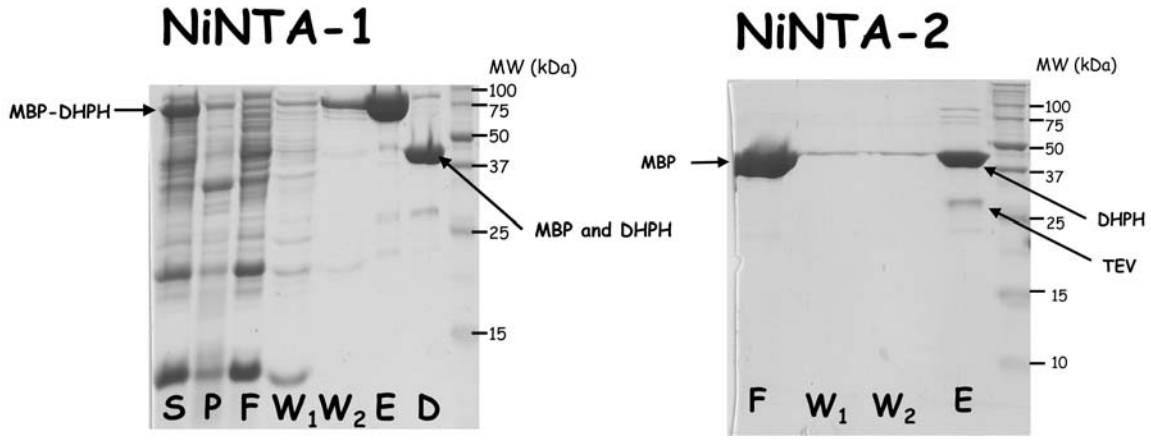

Figure 3.2: Purification of the MBP-DHPH fusion protein and the DHPH domains on Ni-NTA columns. The MBP-DHPH fusion protein was overexpressed in Rosetta (DE3) pLysS cells. Cells were lysed with lysozyme in the presence of protease inhibitors. The cell lysate was applied to a NiNTA column (left panel), the column was washed with 10 column volumes of buffer and the MBP-DHPH fusion protein was eluted with buffer containing 250 mM imidazole. Subsequently, the eluate was dialyzed against buffer without imidazole and in the presence of 1.5 % TEV protease. The digest, which contained MBP (43 kDa) and the DHPH domains (43 kDa), was applied to a second NiNTA column (right panel). The flow through from this column contained MBP and the DHPH domains were eluted as described above. S: supernatant, P: pellet, F: flow through, W_1: wash with lysis buffer, W_2: wash with buffer containing 10 mM imidazole, E: eluate, D: digest.

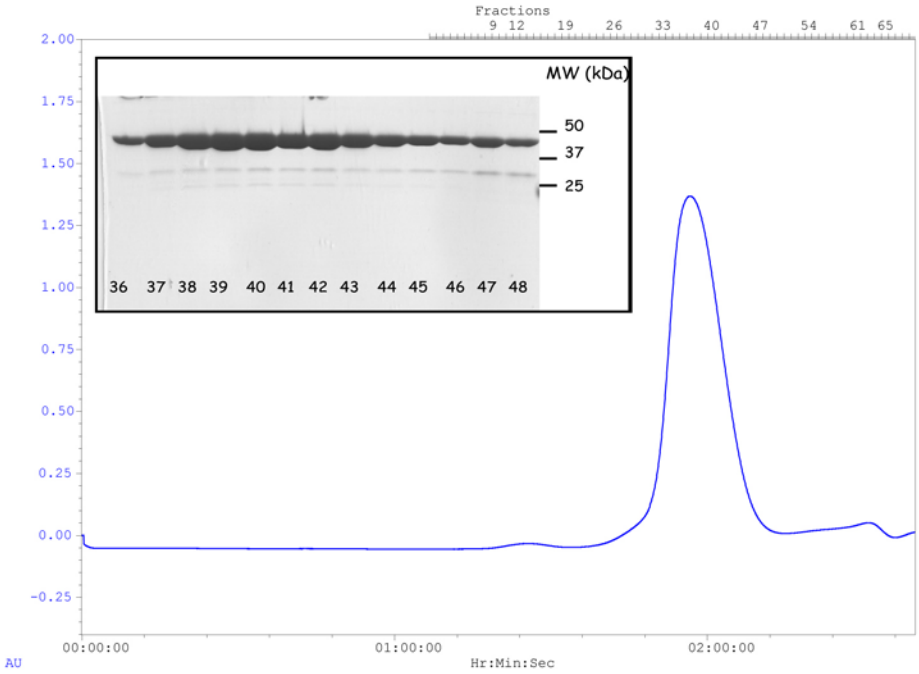

Figure 3.3: Gel filtration of the LARG DH/PH domains. The eluate from the nickel affinity column (NiNTA-2, Fig. 3.2) was concentrated and loaded on a gel filtration column (S200 16/60, Amersham Pharmacia) pre-equilibrated with gel filtration buffer (20 mM HEPES pH 8.0, 10 mM NaCl, 1 mM DTT). The LARG DH/PH domains eluted at a flow rate of 0.75 ml/min and with an apparent MW of 33 kDa. Fractions 36 to 48 were pooled.

TEV protease was expressed from the pRK793 expression vector (a gift from Dr. D. Waugh [93] in the *E.coli* expression strain BL21-RIL (Novagen). Protein expression was induced with 1 mM IPTG and cells were grown at 30° C for 4 hours. Lysis was performed as described above for the DH/PH domains, using TEV-lysis buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 25 mM imidazole pH 8.0, 10 mM βME, 10% glycerol). The lysate was loaded onto a Ni-NTA column (Qiagen) that was pre-equilibrated with TEV-lysis buffer. The column was then washed with 10 column volumes of TEV-lysis

buffer, and TEV protease was eluted in 1 ml fractions with TEV-elution buffer (TEV-lysis buffer with 250 mM imidazole pH 8.0). Fractions containing TEV protease were pooled, concentrated to less than 1 mg/ml, and loaded onto a size exclusion chromatography column (S200 16/60, Amersham Pharmacia Biotech), which was pre-equilibrated with TEV-gel filtration buffer (20 mM HEPES pH 8.0, 200 mM NaCl, 2 mM EDTA, 10 mM DTT, 10% glycerol). Fractions containing TEV protease were pooled, concentrated to approximately 1 mg/ml and stored at -80 ºC.

3.2.2 Crystallization of the LARG DH/PH domains

Initial crystal clusters (Fig 3.4.a) were obtained by the hanging drop vapor diffusion method at 4 ºC from drops containing 1 μl of the concentrated protein (10 mg/ml) and 1 μl of a well solution (15 % PEG 3350, 1 M NaCl and 100 mM sodium citrate pH 6.5) suspended over 1 ml of the well solution. Upon refinement of crystallization conditions, single crystals (Fig. 3.4.b.) were grown at 4 ºC in 8 to 11 % PEG 3350, 0.8 to 1.1 M NaCl and 100 mM sodium citrate pH 6.5. Nucleation occurs within three days and crystals reach their maximum size in about two weeks. Crystals belong to spacegroup *C*2 and have unit cell dimensions of a = 195.5 Å, b = 46.0 Å, c = 75.1 Å and β = 105.0º. The Matthews coefficient (V_M) indicated that there are either 1 or 2 monomers per asymmetric unit (V_M = 3.5 $Å^3$ Da^{-1} and 1.8 $Å^3Da^{-1}$), corresponding to a solvent content of 65 and 31 %, respectively. Structure determination later revealed that the crystals contained only 1 monomer per asymmetric unit. The crystals must be

harvested within several weeks of growth to avoid gradually increasing mosaicity and loss of resolution.

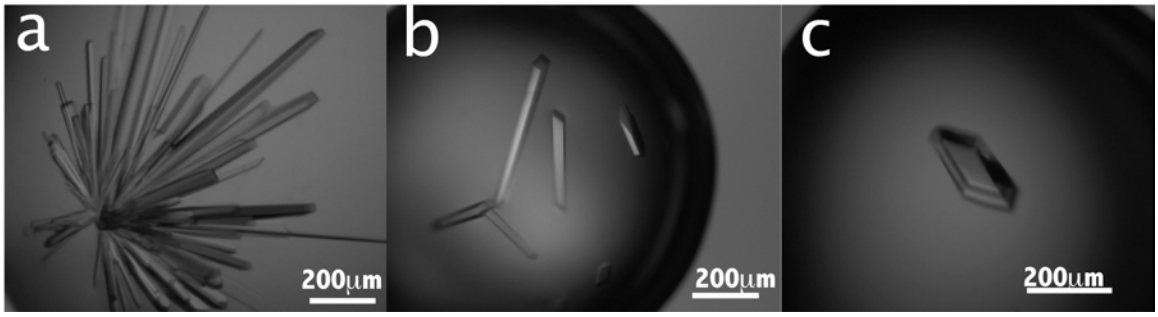

Figure 3.4: Crystals of the DH/PH domains of human Leukemia-associated RhoGEF (LARG). (a) Initial crystals grown as needle clusters (approximate dimensions 0.2 x 0.02 x 0.02 mm). (b) Single crystals grown under optimized conditions. (c) A single crystal grown under optimized conditions (approximate dimensions 0.4 x 0.2 x 0.05 mm).

3.2.3 Harvesting of LARG DH/PH crystals

Crystals were harvested by gradually adding 1 μl cryoprotectant solution at a time to the hanging drop and subsequently transferring crystals to the cyroprotectant solution using cryoloops (Hampton Research). The cryoprotectant solution for the DH/PH domains crystals contained: 15 % ethylene glycol, 20 % PEG3350, 0.6-1 M sodium chloride, 100 mM sodium citrate pH 6.5, 20 mM HEPES pH 8.0, and 2 mM DTT.

3.2.4 Data collection and processing for LARG DH/PH crystals

Native data sets of the DH/PH domains crystal were collected on a MAR345 imaging plate detector using Cu Kα radiation from a Rigaku RU-200H rotating anode

(dhph15, diffraction to 2.5 Å spacings), as well as on a Quantum 4 CCD detector (ADSC) using x-rays from a synchrotron source at BIOCARS beamline 14D at the Advanced Photon Source (APS) (dhph30, dhph45, diffraction to 2.07 and 2.0 Å spacings, respectively) (Tables 3.1 and 3.6). The temperature was maintained at 95 K and 90 K, respectively, with an Oxford Instruments Cryojet. Data sets collected at the APS were used for structure determination and refinement. Data sets of sodium bromide (dhph65) and xenon (dhph63) derivatives of the DH/PH crystals were collected using Cu$K\alpha$ radiation on a MAR345 imaging-plate detector. The temperature was maintained at 95 K with an Oxford Instruments Cryojet.

Data were reduced and scaled using the HKL2000 package [94]. A sigma cut off of -1 was applied. The mosaicity of the diffraction pattern was 0.3.

Crystal	tray	Unit cell dimensions	Resolution (Å)	R_{sym} (%)	R_{merge} (%)	comments
dhph15 (native)	5-21-2, A5	a = 194.3 Å b = 45.9 Å c = 74.9 Å β = 106.2 °	2.5 (home)	5.3		used for initial MR attempts
dhph30 (native)	7-23-2, B3	a = 193.6 Å b = 45.9 Å c = 74.7 Å β = 107.5 °	2.07 (APS)	4.6		used for refinement
dhph45 (native)	8-9-2, D2	a = 195.5 Å b = 46.0 Å c = 75.1 Å β = 105.0 °	2.0 (APS)	4.8		used for MIRAS
dhph63 (Xe) p = 520 psi t = 30min	2-4-3, B4	a = 194.4 Å b = 46.1 Å c = 75.0 Å β = 105.5 °	3.5 (home)	7.9	15.3	used for MIRAS
dhph65 (NaBr) 48 hour soak	2-5-3, B6	a = 194.1 Å b = 45.9 Å c = 74.8 Å β = 106.5 °	2.4 (home)	6.2	15.2	used for MIRAS

Table 3.1: Data collection statistics of crystals used for the structure determination of the LARG DH/PH domains.

Shell Lower limit	Upper limit	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
25.00	4.52	10091.5	272.8	140.0	1.101	0.031	0.036
4.52	3.59	7496.1	250.6	126.5	0.737	0.030	0.033
3.59	3.14	3753.5	167.2	95.6	0.714	0.037	0.036
3.14	2.85	1670.7	104.7	71.8	0.798	0.059	0.054
2.85	2.65	1006.6	87.5	66.2	0.868	0.087	0.075
2.65	2.49	727.8	80.3	66.2	0.967	0.117	0.097
2.49	2.36	554.0	79.0	69.6	1.051	0.162	0.129
2.36	2.26	412.8	79.1	73.4	0.964	0.220	0.165
2.26	2.18	315.8	82.7	78.7	1.010	0.315	0.238
2.18	2.10	239.7	84.2	81.0	0.978	0.413	0.309
All reflections		2675.0	129.8	87.1	0.910	0.047	0.038

Table 3.2: Scaling statistics for the dhph30 crystals. Reflections were integrated with the program DENZO and scaled with the program SCALEPACK.

3.2.5 Structure determination of the LARG DH/PH domains

Molecular replacement is a method for the determination of phases that requires a homologous protein of known structure with approximately 30% sequence identity and structural homology to the protein under scrutiny. Molecular replacement for the phase determination of the LARG DH/PH domains seemed to be the logical first approach for their structure determination because three structures of homologous DH/PH domains (Tiam, Dbs, intersectin) are available and have 23, 25, and 29% sequence identity to the LARG DH/PH domains, respectively. In molecular replacement, a Patterson map of the search model is generated and rotated to fit the Patterson map of the protein under scrutiny. The Patterson function describes the distances between atoms and can be generated without any phase information. Once the orientation of the model is fixed, a translation function is performed, where the search model is moved through the asymmetric unit and structure factors are calculated and compared to the observed structure factors. Peaks from the rotation and translation functions are recorded and

scored. Structure factors can then be generated for the various solutions and, if correct, the resulting phase information can be used to generate initial electron density maps.

The programs MOLREP [95] and EPMR [96] were used and search models were generated from the Tiam, Dbs and intersectin DH/PH domains (PDB codes: 1FOE, 1LB1, 1KI1, respectively). Search models from only the DH domains of these proteins were also used, because the respective PH domains of these proteins assume very different positions relative to the DH domains and it is likely that this is also the case for the LARG DH/PH domains. MR with the DH/PH domains would therefore increase the structural divergence to the search model and decrease the likelihood of finding a correct MR solution. Furthermore, a homology model of LARG, that was based on the closest LARG homolog intersectin, was used as a search model. The homology model was generated with the Swiss pdb-viewer program SWISS-MODEL [97], using intersectin as a template. Poly-alanine models of these DH domains were also generated with the program MOLMAN [98]. The CCP4 suite program MOLREP [95] was used with the option of consideration of the anisotropy of the diffraction pattern. The validity of the molecular replacement solutions was investigated by analyzing crystal packing using the program O [99]. Most molecular replacement solutions had high R factors and low correlation coefficients (Table 3.6) and showed impossible crystal packing, where the DH domains were lined up in sheets and the gaps between sheets were too big to be gapped by the PH domain. Although no convincing translation function solution was found, a recurring rotation function solution was observed when the DH and DH/PH domains of intersectin and the homology model, based on intersectin were used as search models

(Table 3.6). This rotation function solution was later used for the improvement of MIRAS-phased electron density maps (see later section), by translation into density and phase combination with MIRAS-derived phases. Because of the failure of structure determination by MR alone, the two other methods for phase determination, multiple anomalous diffraction (MAD) and multiple isomorphous replacement (MIR), were explored.

Multi-wavelength anomalous diffraction (MAD)

MAD takes advantage of the fact that atoms can absorb x-rays at characteristic wavelengths, which results in changes of the real and an imaginary components of the scattering factor. Diffraction patterns are therefore collected at multiple wavelengths around the absorption edge of the anomalous scatterer to alter the observed structure factor amplitude, and tunable synchrotron radiation is required. Anomalous diffraction can be gained from heavy metal atom derivatives or from proteins that have selenomethionine incorporated instead of methionine, because these elements have accessible absorption edges. The LARG DH/PH domains contain 6 methionines, which would allow a MAD structure determination.

Production of Selenomethionine substituted LARG DH/PH domains

A protocol for the incorporation of selenomethionine into recombinant protein by LeMaster was followed [100]. The methionine auxotroph *E. coli* strain B384 (DE3) pLysS was used as the expression strain for the LARG DH/PH domains. A starter culture of 50 ml LB media was inoculated, incubated overnight and expanded to 1 liter of

LeMaster selenomethionine medium, which contains glucose, 17 amino acids, the bases for nucleic acids, various salts, sulfate, and 50 mg selenomethionine per liter of medium. The culture was grown to OD_{600} = 0.6 and harvested and purified as described for the native LARG DH/PH domains (3.2.1).

Although selenomethionine-incorporated LARG DH/PH domains were successfully produced and crystallized in the form of needle clusters, refinement of the crystallization conditions did not result in well diffracting crystals. Only long, thin needles (approximate dimensions: 0.2 x 0.02 x 0.02 mm) could be harvested from the needle clusters, and they showed only weak diffraction. Furhtermore, these clusters did not exhibit the characteristic selenium absorption edge in a fluorescence scan at the synchrotron.

Multiple isomorphous replacement with anomalous scattering (MIRAS)

The structure of the LARG DH/PH domains was finally determined by a combination of multiple isomorphous replacement with anomalous scattering (MIRAS) and molecular replacement (MR). Multiple isomorphous replacement requires the derivatization of the protein crystals by a strong scatterer, preferably a heavy metal compound. The positions of the heavy metal atoms can be determined by inspection of a difference Patterson map. Difference Patterson maps are generated by a Fourier synthesis with the phaseless quantitiy $\overline{|F|^2}$ (structure factor squared) as the coefficient instead of $\overline{F}$. Peaks in the Patterson maps are distance vectors between atoms. Peaks corresponging to difference vectors between symmetry related heavy atoms are

accumulated at special sections, called the Harker sections. Cross peaks between heavy atoms can be found on non-Harker sections. When two heavy atom positions have the same y-coordinate, in spacegroup C2, their cross peaks are also found on Harker sections. Once the coordinates of the Patterson peaks are known the heavy metal positions in the unit cell and their phase angle (α_H) and structure factors (F_H) can be determined. The protein phase angles (α_P) are then derived from the structure factors of the protein (F_P) which can be derived from $F_P = F_{PH}-F_H$. The phase information can then be used to generate an electron density map.

Derivatization of the DH/PH domains crystals was initially attempted with several platinum and mercury compounds by soaking the crystals in solutions of these compounds at varying concentrations for varying times. Soaking in thimerosal at 0.5 mM for 1 hr lead to cracks in the crystals. Soaks in 0.5 mM K_2PtCl_4, platinum pyridinium chloride, and mercury fluorescein preserved isomorphism and diffraction (to 2.5 Å and 3.2 Å, respectively) but did not result in useful derivatives as determined by scaling with HKL2000 [94] or inspecting difference density maps.

Many heavy atom compounds have useful anomalous signals at the Cu$K\alpha$ wavelength and anomalous scattering can be included in isomorphous phasing.
A xenon derivative was generated by subjecting a crystal to 520 psi xenon in a xenon chamber (Hampton Research) for 30 min at 4 ˚C. A bromide derivative was generated by soaking crystals in cryoprotectant solution, with NaBr substituting for NaCl, for 48 hours. Data collection and processing was done as described above for the native DH/PH domains, except that a full 360º of data was collected. This is to ensure all Bijovet pairs

are collected ata reasonably high redundancy. Collecting as many degrees as possible improves the statistics. The xenon and NaBr derivatives were assessed for their potential as isomorphous derivatives by scaling derivative and native data sets together using the program SCALEPACK of the *HKL* package [94]. Two parameters in the scaling output file, χ^2 and R_{sym}, give information about the isomorphism of the two crystals and the quality of the derivative. χ^2 is a measure of the error associated with fitting an observed reflection to a predicted reflection and should have a value of 1-2 if the sigmas of I (intensities) are estimated correctly. If a native data set and that of a derivative are scaled together, χ^2 will ideally have values of around 10 or higher. If χ^2 is ~1 no useful derivative is present. If χ^2 is very large (~50) the derivative crystal is probably non-isomorphous. The second parameter in the scaling output file that can give information about the presence of a derivative is R_{merge}, which is R_{sym} of the merged data sets. R_{sym} is a measure of the errors within a data set and more specifically of the agreement of the intensities of symmetry related reflections. Two independent data sets will have a theoretical R_{merge} of 0.57. R_{sym} typically increases with increasing resolution, because high resolution data cannot be measured accurately. When a derivative and native data set are scaled together and they are isomorphous, the R_{merge} distribution with increased resolution will ideally be flat. An increasing distribution is indicative of non-isomorphous crystals. The xenon and sodim bromide derivatives had a higher degree of isomorphism with different native datasets (dhph45 and dhph30, respectively; see Table 3.1). Because, the phasing statistics for the xenon derivative were better, the dhph45 data set was used for MIRAS phasing (Table 3.1). However, the other native data set

(dhph30) was used for refinement, because it was more complete (98.3 % for dhph30 versus 96.8 % for dhph45).

After evaluation of the derivatives by scaling in SCALEPACK (Tables 3.3 and 3.4), derivative and native data sets were merged using the program CAD from the CCP4 program suite [101] and scaled together with the program SCALEIT [101]. Subsequently, isomorphous and anomalous difference Patterson maps were generated. The xenon derivative generated interpretable isomorphous and anomalous difference Patterson maps.

Three prominent peaks were seen on the Harker sections as well as one prominent peak on a section other than the Harker section (Figure 3.5). The same peak pattern was observed in isomorphous and anomalous difference Patterson maps (Figure 3.7). One of the prominent peaks and a weaker peak on the Harker section were identified as xenon site 2 (Xe2) and xenon site 1 (Xe1). The other two prominent peaks were identified as cross peaks of two sets of symmetry related Xe1 and Xe 2. Both cross peaks occur on the Harker section because Xe1 and Xe2 have the same y coordinates. The two peaks on the Harker section were refined using the program MLPHARE [102]. A third xenon site (Xe 3) was identified with the help of difference Fourier maps and the cross peaks on section v = 0.12 (Figure 3.6).

Shell	Lower limit	Upper limit	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
	20.00	6.40	18882.6	499.2	499.2	43.905	0.154	0.213
	6.40	5.11	8292.7	219.3	219.3	18.756	0.142	0.193
	5.11	4.47	13307.2	344.6	344.6	15.390	0.125	0.189
	4.47	4.06	11890.4	292.1	292.1	14.750	0.132	0.190
	4.06	3.78	9191.4	240.0	240.0	11.365	0.149	0.219
	3.78	3.55	7937.2	229.7	229.7	6.983	0.143	0.194
	3.55	3.38	6155.0	213.1	213.1	6.044	0.178	0.264
	3.38	3.23	4438.5	195.1	195.1	3.594	0.184	0.212
	3.23	3.11	3216.8	164.7	164.7	2.810	0.224	0.283
	3.11	3.00	2365.6	138.2	138.2	1.855	0.241	0.292
All reflections			8571.3	253.6	253.6	11.711	0.153	0.207

Table 3.3: Scaling statistics for the combined scaling of native (dhph45) and xenon derivative (dhph63) data sets. Reflections were integrated with the program DENZO and scaled with the program SCALEPACK.

Shell	Lower limit	Upper limit	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
	99.00	4.76	14302.4	262.4	262.4	170.157	0.129	0.148
	4.76	3.78	10862.9	207.3	207.3	114.459	0.128	0.143
	3.78	3.30	6345.4	139.3	139.3	83.025	0.158	0.179
	3.30	3.00	2915.8	97.5	97.5	40.257	0.177	0.223
	3.00	2.78	1614.5	79.1	79.1	21.587	0.203	0.268
	2.78	2.62	1100.7	77.7	77.7	11.569	0.216	0.262
	2.62	2.49	842.7	77.6	77.6	6.777	0.237	0.299
	2.49	2.38	643.8	72.5	72.5	5.713	0.273	0.337
	2.38	2.29	509.9	70.9	70.9	5.208	0.317	0.385
	2.29	2.21	444.5	75.8	75.8	6.677	0.355	0.907
All reflections			3925.9	115.6	115.6	47.509	0.152	0.158

Table 3.4: Scaling statistics for the combined scaling of native (dhph45) and NaBr derivative (dhph65) data sets. Reflections were integrated with the program DENZO and scaled with the program SCALEPACK.

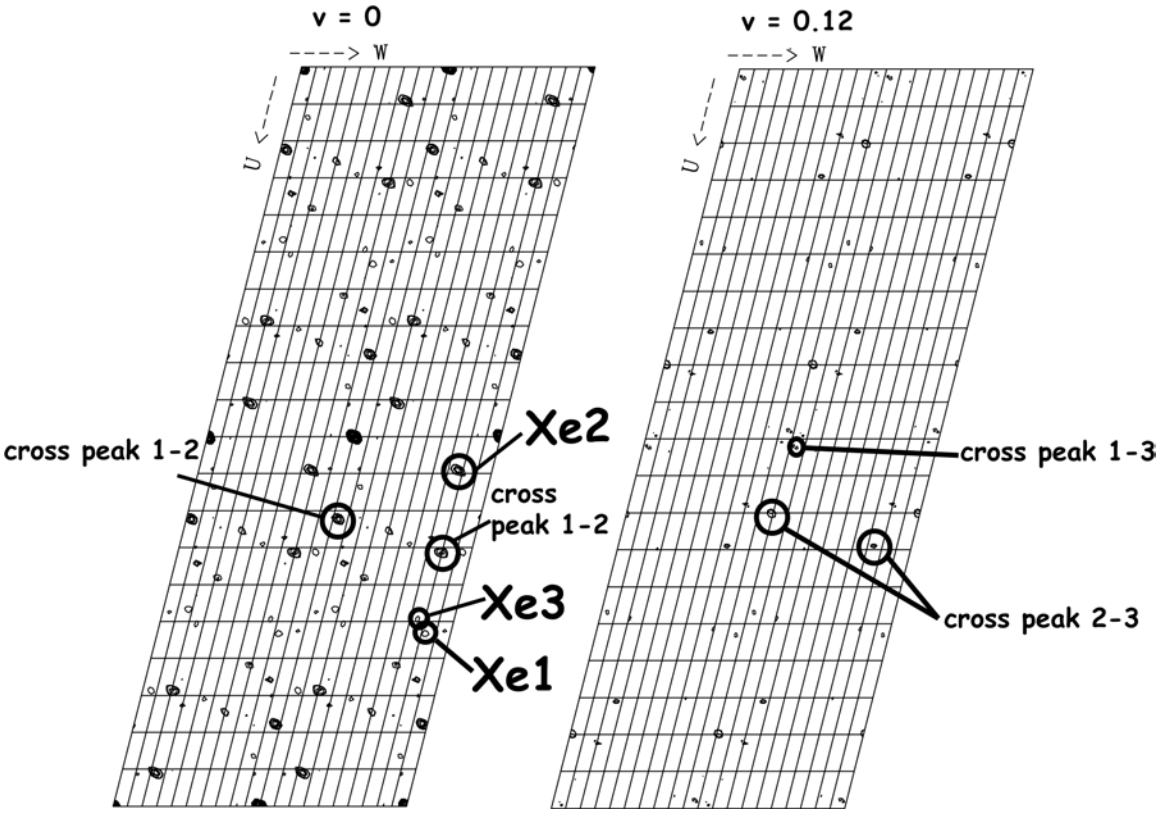

Figure 3.5: Patterson maps of the xenon derivative of the LARG DH/PH domains. Four unit cells of the Harker section (v = 0) and section v = 0.12 of the isomorphous difference Patterson map are shown, contoured starting at 2.0 σ in steps of 4.0 σ for the Harker section and in steps of 2.0 σ for the section v = 0.12. Xenon peaks and cross peaks of one asymmetric unit are circled

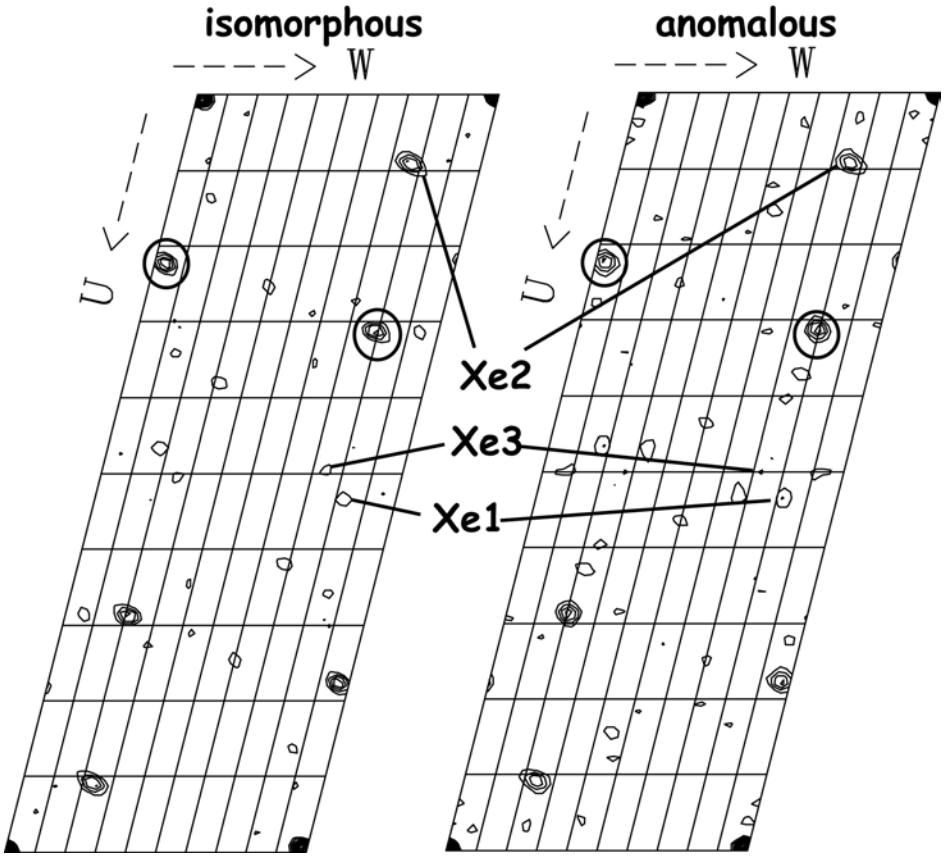

Figure 3.6: Isomorphous and anomalous difference Patterson maps show the same peaks. One unit cell of the Harker section (v = 0) of isomorphous and anomalous difference Patterson maps of the DH/PH xenon derivative contoured starting at 1.0 σ in steps of 1.0 σ are shown.

The NaBr derivative had only one site, which was identified in difference Fourier maps generated using the xenon-derived phases. The statistics for the xenon and bromide site refinement are shown in Table 3.5. Although, the phasing power for the sodium bromide derivative is unusually low, refinement statistics (*i.e.* lack of closure, figure of merit) for the bromide site in MLPHARE indicated that this derivative contributed to the phase information. To further investigate the contribution of the individual derivatives to

the phase, a plot of the cosine of the phase differences against resolution was generated for the xenon and sodium bromide derivatives (Figure 3.7). This plot shows that the anomalous phase information of the xenon derivative and the phase information of the bromide derivative only contributed to the correct phases at high resolution (>3.8 Å for the xenon anomalous phases and >3.1 Å for the bromide isomorphous phases).

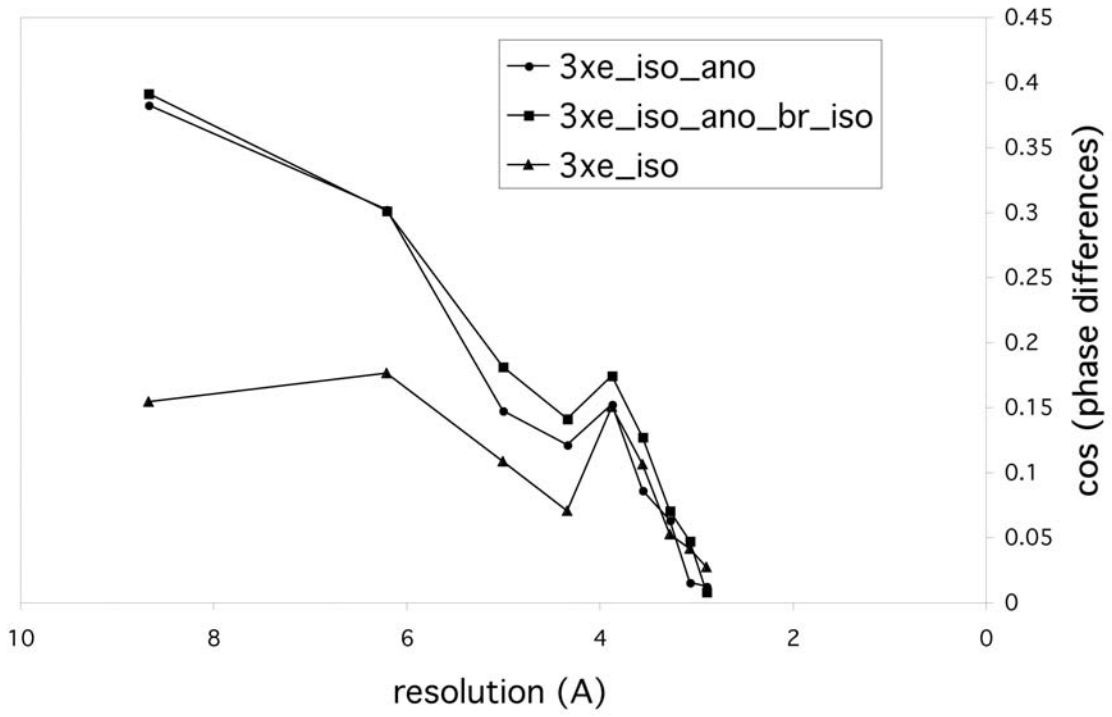

Figure 3.7: The anomalous signal of the xenon derivative and the isomorphous signal of the bromide derivative contribute to the phase information only at high resolution. The resolution-dependent mean phase discrepancies for the xenon and bromide derivative phases are shown.

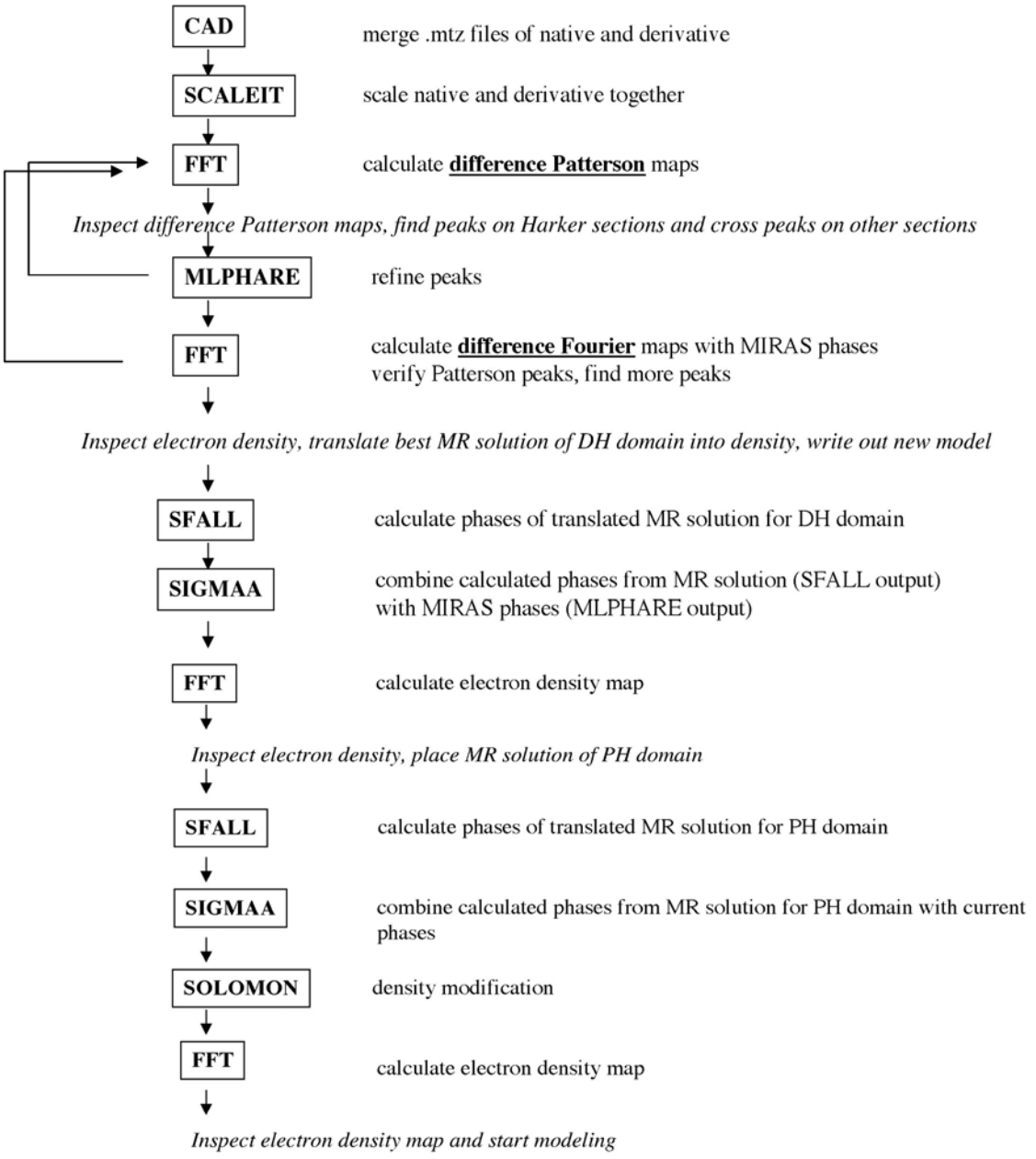

Figure 3.8: Flow chart of the steps involved in the phase determination of the LARG DH/PH domains by a combination of MIRAS and MR.

Molecular Replacement (MR)

Inspection of the MIRAS-phased electron density maps in the program O [99] made it possible to manually place a molecular replacement solution of a homology model of the LARG DH domain (see 3.2.9). The molecular replacement solution was generated with the program MOLREP of the CCP4 suite [95] and had a peak height (RF/σ) of 6.51. The manual placement of the molecular replacement solution only involved translation and not rotation. Phases of the manually placed molecular replacement solution were calculated with the program SFALL [101] and combined with MIRAS-derived phases from the MLPHARE output using the program SIGMAA [101] (see Fig. 3.8). After phase combination the electron density improved considerably and the resulting electron density maps allowed placement of a homology model of the LARG PH domain. The quality of the electron density further improved after phase combination of the PH and DH domains and density modification with the program SOLOMON from the CCP4 suite [101]. Model building in O [99] was then alternated initially with refinement in CNS [103] and subsequently in REFMAC including TLS refinement (see 3.2.11). However, electron density remained ambiguous for the N-terminus of the DH domain (residues 765-770), the N-terminus of the PH domain (999-1008), the β1-β2 loop (1029-1035), the insert in the β4 strand (1062-1073), the β5-β6 loop (1094, 1095), and the C-terminus of the PH domain (1136-1138) (Figure 3.15). Model building was further assisted by the structure of the DH/PH-RhoA complex (see next section), which facilitated interpretation of these regions and the final structure of the DH/PH domains lacks one residue at the N-terminus (residue 765) of the DH domain

as well as seven residues within the βN-$N_{3/10}$ loop (residues 1000-1006), 13 residues (residues 1062-1074) within the β4 strand (1062-1074), and one residue (residue 1093) within the β5/β6 loop of the PH domain. Ninetytwo water molecules were included in the structure and only assigned to the DH domain. The PH domain has considerably higher B-factors and no obvious water molecules were evident in the electron density maps. The criteria used for the water molecules were: (1) the peak in the electron density map is at least 3σ, (2) the water molecule is involved in at least two hydrogen bonds, and (3) the hydrogen bond distance does not exceed 3.5 Å.

3.2.6 Refinement of the LARG DH/PH structure

Initial simulated annealing and individual B-factor refinement was performed in CNS [103] followed by refinement with the program REFMAC [101] with TLS refinement [104]. TLS refinement is a method that takes into account the displacement of atoms due to anisotropy of the data. In TLS refinement it is assumed that the displacement of neighboring atoms is correlated. Therefore pseudo-rigid bodies are defined (TLS groups), for which refinement parameters are selected, that describe the **t**ranslation, **l**ibration, and **s**crew-rotation of the pseudo-rigid bodies. The LARG DH domain and the PH domain were considered as separate TLS groups. Waters were only assigned to the DH domain and therefore grouped with the DH domain. The structure was refined to final R_{work} and R_{free} of 23.1 and 27.6%, respectively (Table 3.6).

	dhph30 native **refinement**	**dhph45** native **MIRAS phasing**	**dhph63** Xenon	**dhph65** NaBr	**rhoAdhph92** **MR phasing and refinement**
Wavelength (Å)	1.0	0.9794	1.5418	1.5418	1.0688
Spacegroup	*C*2	*C*2	*C*2	*C*2	*C*2
Unit cell (Å, º)	a = 193.6	a = 195.5	a = 194.4	a = 194.1	a = 294.91
	b = 45.9	b = 46.0	b = 46.1	b = 45.9	b = 95.0
	c = 74.7	c = 75.13	c = 75.0	c = 74.8	c = 157.03
	β = 107.5	β = 105.0	β = 104.5	β = 106.6	β = 94.04
Resolution (Å)	2.13	2.05	3.5	2.4	3.2
Unique reflections (total)	35477 (814907)	41093 (518961)	8384 (67240)	25235 (120925)	70549 (1360421)
Completeness (%)[†]	98.3 (93.7)	96.8 (97.8)	100 (100)	99.8 (100)	91.2 (57.4)
R_{sym} (%) [§]	4.6 (36.5)	4.8 (51.4)	7.9 (16.1)	6.2 (42.3)	8.3 (57.4)
Average $I/\sigma(I)$	21.0 (3.3)	24.0 (3.1)	28.6 (12.6)	32.2 (4.8)	11.4 (1.4)
Phasing power[‡]					
Cenctric	-	-	0.98	0.19	-
Acentric	-	-	1.19	0.26	-
Resolution range for refinement (Å)	24 to 2.07	-	-	-	158.11 to 3.19
Total reflections used	35784	-	-	-	61716
Number of protein atoms	2880	-	-	-	16906
Number of water molecules	92	-	-	-	-
RMSD bond lengths (Å)	0.014	-	-	-	0.04
RMSD bond angles (°)	1.31	-	-	-	2.91
R_{work}[‖] (%)	23.8	-	-	-	25.6
R_{free}[¶] (%)	28.2	-	-	-	30.6
Average B-factor (Å^2)	34.9	-	-	-	45.0

[†]Values in parentheses refer to the highest resolution shell (DH/PH: 2.12-2.05 Å, DH/PH-RhoA: 3.31-3.2 Å).
[§]$R_{sym} = \Sigma |I-I_{avg}|/\Sigma I$, where the summation is over all symmetry-equivalent reflections, excluding reflections observed only once.
[‖] $R_{work} = \sum_h ||\mathbf{F_{obs}}(h)|-|\mathbf{F_{calc}}(h)|| / \sum_h |\mathbf{F_{obs}}(h)|$; no I/σ cutoff was used during refinement.
[¶] 5% of reflections were reserved from refinement for the calculation of R_{free}.
‡ Phasing power = (F_H/LOC), where LOC is the lack of closure.

Table 3.5: Data collection, phasing and refinement statistics for the LARG DH/PH domains and the DH/PH-RhoA complex.

Stereochemistry

Two residues in the DH/PH domains, Ser^{833} and Asp^{1054}, fall within the disallowed region of the Ramachandran plot. In other atomic structures of DH domains, residues equivalent to Ser^{833} have the same strained backbone stereochemistry suggesting its conformation is required to maintain the DH domain fold. Ser^{833} in the LARG DH domain seems to be distorted because the adjacent Phe^{832}, which is highly conserved in RhoGEF DH domains, participates in a conserved hydrophobic core that holds together helices α1, α3, α4, and α5. Thus, the DH domain can energetically afford to keep a residue with disallowed stereochemistry because it gains binding energy through interactions in the hydrophobic core. Asp^{1054} exists in the i+1 position of a type I β-turn in the PH domain (β3-β4 loop), which in most turns is normally occupied by glycine.

3.2.7 Analysis of crystal packing in the DH/PH crystals

Analysis of the packing of DH/PH molecules reveals a hydrophobic patch on the PH domain that could potentially be involved in protein-protein interactions and exploited for regulation of the DH/PH domains. PH domains in the DH/PH domains crystals interact via a hydrophobic interface formed by the β4-β7 sheet, involving exposed aromatic and aliphatic residues (Leu^{1086}, Phe^{1098}, Ile^{1109}, Ile^{1100}, Ala^{1107}). Interestingly, the same site on the PH domain is involved in crystal contacts of the DH/PH-RhoA complex, further confirming that this hydrophobic patch is a potential protein-docking site (see 3.3.10).

Analysis of the crystal packing of the DH/PH domains crystals also explains the high degree of anisotropy in these crystals. The diffraction pattern showed good diffraction along the z-axis of the crystals and poor diffraction along the x-axis of the crystals. This is a result of alignment of the DH domains in sheets that run parallel to the y-z plane, and connection of the sheets by the highly mobile PH domains that align parallel to the x-axis. This formation leaves the crystal lattice with channels that measure about 70 x 60 Å in the x and z directions, respectively.

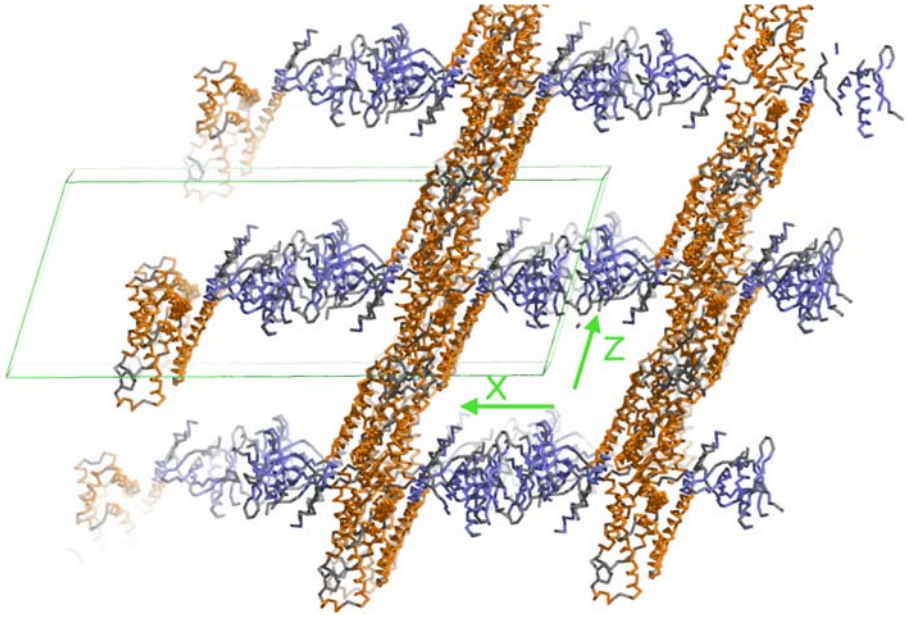

Figure 3.9: Crystal packing of the LARG DH/PH domains in spacegroup C2, looking down the y-axis. The DH and PH domains are rendered as orange and blue tubes, respectively. One unit cell is shown as a green box. DH domains aligned in sheets that run parallel to the yz plane. PH domains align parallel to the xy plane. The crystal packing explains the good diffraction along the z-axis but poor diffraction along the x-axis.

3.2.8 Analysis of the failure of molecular replacement

Molecular replacement for the phase determination of the LARG DH/PH domains seemed to be the logical first approach for their structure determination since three structures of homologous DH/PH domains (Tiam, Dbs, Intersectin) are available and have 23, 25, and 29% sequence identity to the LARG DH/PH domain, respectively. However, solving the phase solely by molecular replacement was not possible with the progams MOLREP [95] and EPMR [96]. This is not surprising if the DH/PH domains are used as a search model, because the relative disposition of DH to PH domain is highly divergent in RhoGEFs (Fig. 3.19). However, RhoGEF DH domains are structurally fairly homologous and it was therefore surprising that even when the DH domain by itself was used as a search model it did not yield the correct molecular replacement solution. The CCP4 suite program MOLREP [101] was run with the option of consideration of the anisotropy of the diffraction pattern and it resulted in a recurring rotation function solution when intersectin and the homology model were used as search models (Table 3.6). However, no convincing translation function peaks were found. The best molecular replacement solutions were visualized in the unit cell using the program O [99]. But the packing of the DH domain usually showed overlaps or the distances between DH domains were to big to be gapped by PH domains and the solutions would therefore not be consistent with a protein lattice. Since the DH domains of RhoGEFs whose structures have been previously determined (Tiam, Intersectin, Dbs, Vav, β-PIX, Trio) are all structurally very similar and PH domains also share close structural homology it was concluded that one reason for the failure of the molecular replacement programs to pick a

convincing solutions was likely the arrangement of the PH domain relative to the DH domain. Indeed the realtive arrangement of the LARG DH and PH domains is different from that of intersectin. Although the arrangement in LARG is similar to that of Dbs, Dbs was also not a good search model, possibly due to a 13 Å translation of the PH domain along the α6 segment of the Dbs DH domain.

In an attempt to analyze the failure of molecular replacement, the solved structure of the DH/PH domains was used to search for a MR solution with intersectin or the homology model, based on intersectin as the search models (Table 3.6). However, inspection of the best solutions in the program O revealed that MOLREP could not pick the correct packing of the DH domains within the layers of DH domains that are present in the crystals (see Fig. 3.9).

MR_search_model	#solution	alpha	beta	gamma	Xfrac	Yfrac	Zfrac	TF/sig	R-fac	Corr
final_model_DH	*Sol_Tf_1*	*0.00*	*0.00*	*0.00*	*0.235*	*0.000*	*0.227*		*0.455*	*0.475*
homologymod_DH	Sol_TF_1	221.02	65.78	155.06	0.235	0.000	0.196	7.63	0.589	0.227
Intersectin_DH	Sol_TF_1	220.69	68.28	154.13	0.238	0.000	0.223	6.97	0.580	0.237
final_model_DHPH	*Sol_Tf_1*	*0.00*	*0.00*	*0.00*	*0.180*	*0.000*	*0.325*		*0.436*	*0.520*
homologymod_DHPH	Sol_TF_1	221.67	65.56	153.82	0.176	0.000	0.308	5.80	0.586	0.231
Intersectin_DHPH	Sol_TF_1	221.30	67.89	153.99	0.183	0.000	0.305	5.58	0.597	0.212

Table 3.6: Molecular replacement (MR) using the final model of the LARG DH and DH/PH domains as a search model, in comparison with MR solutions using the DH and DH/PH domains of a homology model and intersectin as the search models. The CCP4 program molrep picked a recurring rotation function solution. Although some translation function peaks of intersectin and the homology model seem to be equivalent to the peaks of the final model, the fractional coordinates cannot be directly compared because MOLREP moves each model to its center of mass, which is different for each model. However, inspection of the solutions in the program O revealed that the molrep solutions occupied the same space as the final model in the unit cell but that MOLREP could not pick the correct packing of the DH domains within the layers of DH domains that are present in the crystals (see Fig. 3.9).

3.2.9 Expression and Purification of RhoA in *E.coli* from pGEX-KG-RhoA

Human RhoA (residues 1-193) was expressed from the pGEX-KG-RhoA vector (a gift from Dr. T. Kozasa) as an N-terminal fusion protein with **g**lutatione-**S**-**t**ransferase (GST) containing a thrombin cleavage site. Rosetta (DE3) pLysS cells (Novagen), harboring the pGEX-KG-RhoA plasmid were grown to OD_{600} = 0.6, induced with 1 mM IPTG at 30 °C and harvested after 5 hrs by centrifugation at 3500 rpm in a Beckman Coulter Avanti J-20 centrifuge, using a JLA 8.100 rotor. The cell pellet was frozen in liquid nitrogen and stored at -80 ºC until purification.

Cells were lysed with lysozyme (Sigma) (1 mg/ml) in lysis buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 10 % Glycerol, 5 mM DTT, 50 μM GDP) in the presence of 0.3 mM EDTA and protease inhibitors (1 μM leupeptin, 1 mM lima bean trypsin inhibitor, 1 mM PMSF, and 1 mM TPCK). Lysates were supplemented with 10 mM $MgCl_2$ and 20 μg DNaseI (Sigma) per gram of *E.coli* and centrifuged for 45 min at 35000 rpm in a Beckman Coulter Optima LE-80K ultracentrifuge, using a Type 45 Ti rotor. The supernatant was loaded onto a drip column containing glutathione-sepharose 4B (Amersham Pharmacia Biotech Inc), which was pre-equilibrated with equilibration buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 10 % glycerol, 1 mM DTT, 50 μM GDP). After washing with 10 column volumes of equilibration buffer, protein was eluted in 1 ml fractions in equilibration buffer supplemented with 10 mM glutathione. Fractions containing the GST-RhoA fusion protein were pooled and dialysed overnight against dialysis buffer (20 mM HEPES pH

8.0, 150 mM NaCl, 10 mM $MgCl_2$, 10 % glycerol, 1 mM DTT, 5 μM GDP, 2.5 mM $CaCl_2$) in the presence of 1.7 % thrombin. GST and undigested GST-RhoA fusion protein were separated from RhoA via a glutathione-sepharose 4B column. The flow-through, which contained RhoA was concentrated, and loaded onto a size exclusion chromatography column (S200 16/60, Amersham Pharmacia Biotech) pre-equilibrated with gel filtration buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1mM $MgCl_2$, 40 μM GDP, 1 mM DTT). Fractions containing RhoA were pooled, concentrated to approximately 5 mg/ml and stored at –80 ºC. The typical yield was 2.5 mg per liter of cell culture.

3.2.10 Cloning, expression and purification of RhoA from a pMAL vector

RhoA expressed from the pGEX-KG-RhoA plasmid contains 15 additional amino acids at the N-terminus. To eliminate the possibility that these additional amino acids interfere with crystallization, RhoA was cloned from the pGEX-KG-RhoA vector (see 3.2.2) into the modified pMAL expression vector pMALc2H_{10}T, which encodes an N-terminal MBP affinity tag followed by an H_{10} tag and TEV protease cleavage site (pMALc2H_{10}T_RhoA, Figure 3.2.). The pMALc2H_{10}T vector was generated by inserting 5' phosphorylated oligonucleotides (tcg ggc acc atc acc atc acc atc acc atc acc atg aaa acc tgt att ttc agg gag, aat tct ccc tga aaa tac agg ttt tca tgg tga tgg tga tgg tga tgg tga tgg tgc, Integrated DNA technologies) that encode the H_{10} tag and TEV protease cleavage site between the restriction sites *Ava*I and *Eco*RI of the pMALc2X vector (NEB).

RhoA was expressed from the pMALc2H_{10}T_RhoA vector as an N-terminal fusion protein with MBP and an N-terminal hexahistidine tag. Both, MBP and the H_{10} tag can be cleaved with TEV protease. Rosetta (DE3) pLysS cells (Novagen) carrying the pMALc2H_{10}T plasmid were grown to OD_{600} = 0.6, induced with 1 mM IPTG at 30 ºC and harvested after 4-6 hrs by centrifugation at 3500 rpm in a Beckman Coulter Avanti J-20 centrifuge, using a JLA 8.100 rotor. The cell pellet was frozen in liquid nitrogen and stored at -80 ºC until purification.

Lysis was done as described above (3.2.2) for RhoA expressed from the pGEX-KG-RhoA vector except that lysis buffer contained 10 mM βME instead of DTT. The supernatant was loaded onto a drip column containing Ni-NTA resin (Qiagen), which was pre-equilibrated with equilibration buffer (20 mM HEPES pH 8.0, 300 mM NaCl, 10 mM Imidazole, 10 mM βME, 10 % glycerol, 10 mM $MgCl_2$, 5 μM GDP). After washing with 5 column volumes of equilibration buffer and 5 column volumes of wash buffer (equilibration buffer with 20 mM imidazole) protein was eluted in 1 ml fractions in elution buffer (equilibration buffer with 250 mM imidazole). Fractions containing the MBP-RhoA fusion protein were pooled and dialyzed overnight against dialysis buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 5 mM βME, 10 % Glycerol, 5 mM $MgCl_2$, 5 μM GDP) in the presence of 2 % TEV protease. RhoA was separated from MBP and undigested MBP-RhoA fusion protein by subjecting the dialyzed solution to another Ni-NTA column. The flow-through, which contained RhoA, was concentrated and loaded onto a size exclusion chromatography column as described in 3.2.2. Fractions containing

RhoA were pooled, concentrated to approximately 5 mg/ml and stored at –80 ºC. RhoA expressed from both, the pGEX-KG-RhoA and the pMALc2H$_{10}$T_RhoA vectors showed smeared bands on SDS-PAGE gels (Figure 3.3b). This phenomenon has been observed previously for Ras and Rho family GTPases and was thought to be due to oxidation of the GTPases [105].

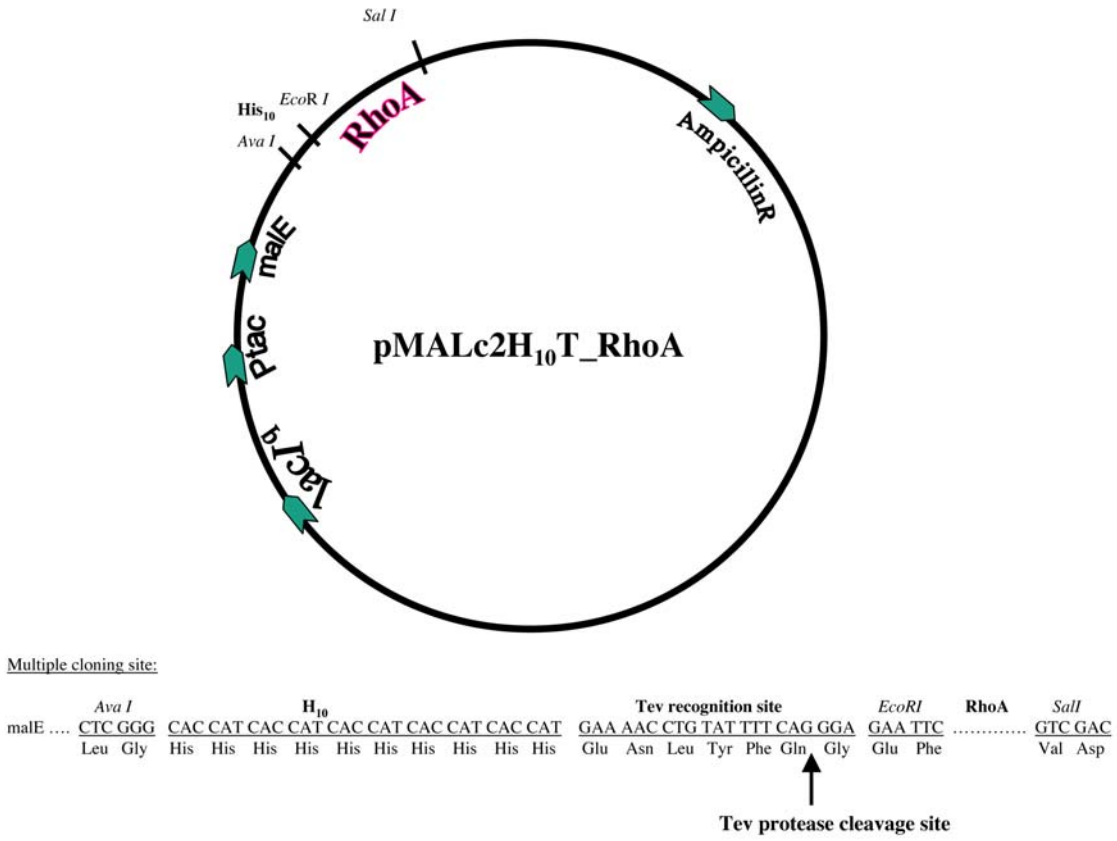

Figure 3.10: Vector map of the expression vector for RhoA, pMALc2H$_{10}$T_RhoA. RhoA was expressed as N-terminal fusion protein with maltose binding protein (MBP) followed by an H$_{10}$ tag and the TEV protease cleavage site. After cleavage with TEV protease RhoA contains additional three amino acids at the N-terminus. Arrows indicate the start of the coding regions for the Ptac promotor (Ptac), lac repressor (lacI^q), β-lactamase (AmpicillinR), and maltose binding protein (malE).

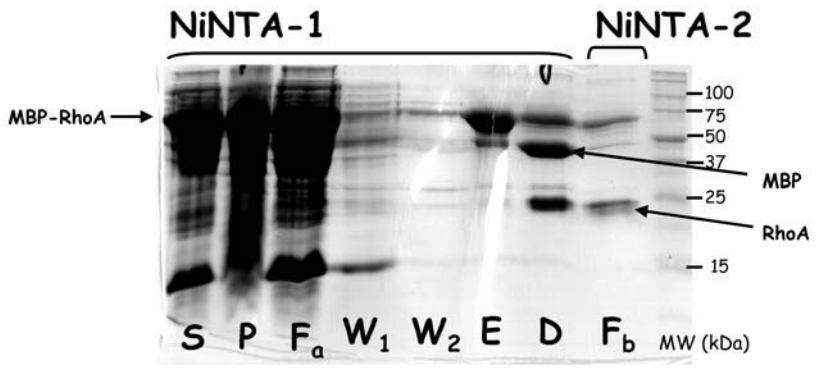

Figure 3.11: Purification of RhoA using NiNTA columns. The MBP-RhoA fusion protein was overexpressed in Rosetta (DE3) pLysS cells. Cells were lysed with lysozyme in the presence of protease inhibitors. The cell lysate was applied to a NiNTA column (NiNTA-1), the column was washed with 10 column volumes of buffer and the MBP-RhoA fusion protein was eluted with buffer containing 250 mM imidazole. Subsequently, the eluate was dialyzed against buffer without imidazole and in the presence of 2 % TEV protease. The digest, which contained MBP (43 kDa) and RhoA (21 kDa), was applied to a second NiNTA column (NiNTA-2). The flow through from this column (Fb) contained RhoA and undigested MBP-RhoA fusion protein. S: supernatant, P: pellet, Fa: flow through from NiNTA-1, W_1: wash with lysis buffer, W_2: wash with buffer containing 10 mM imidazole, E: eluate, D: digest, Fb: flow through from NiNTA-2.

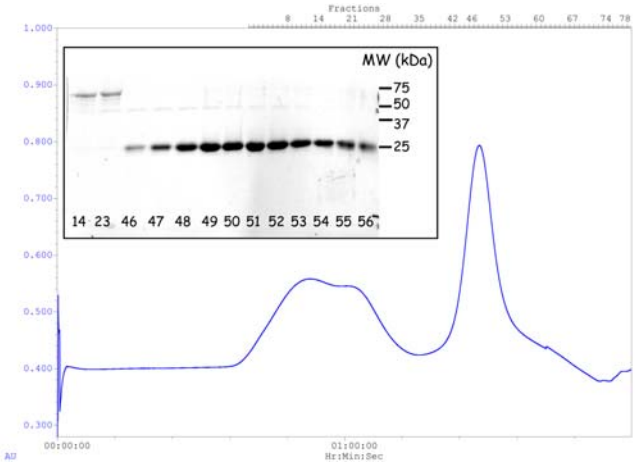

Figure 3.12: Gel filtration of RhoA. The flow through of the NiNTA-2 column was concentrated and applied to a gel filtration column (S200 16/60, Amersham Pharmacia Biotech) pre-equilibrated with gel filtration buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM $MgCl_2$, 40 μM GDP, 1 mM DTT). RhoA eluted at a flow rate of 1 ml/min and with an apparent MW of 19 kDa. Fractions 46 to 57 were pooled. Unidgested MBP-RhoA and MBP eluted as a broad peak.

3.2.11 Purification of the RhoA:DH/PH complex

The DH/PH domains of LARG were mixed with a two-fold molar excess of RhoA and diluted 10-fold with complex buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 10 mM EDTA, 2 mM DTT). After incubation for 10 min on ice, the complex was loaded onto a size exclusion chromatography column (S200 16/60, Amersham Pharmacia Biotech), which was pre-equilibrated with gel filtration buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM EDTA, 2 mM DTT). Fractions containing the RhoA-DH/PH complex were pooled, concentrated to approximately 10 mg/ml and stored at -80 °C until crystallization.

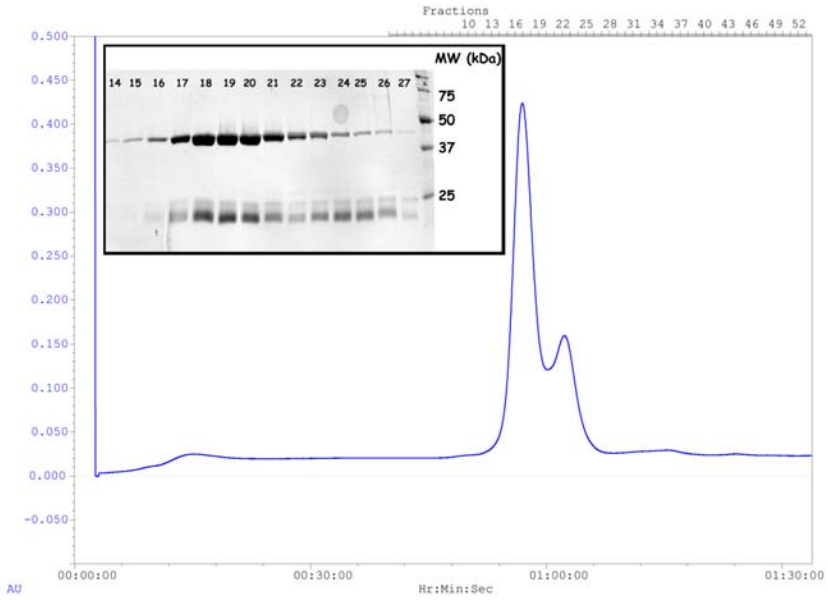

Figure 3.13: Chromatogram of the gel filtration of the RhoA-DH/PH complex. The DH/PH domains were mixed with a two-fold molar excess of RhoA in buffer containing 10 mM EDTA and incubated on ice for 10 min. Subsequently the sample was loaded onto a gel filtration column and the protein was eluted at a flow rate of 0.5 ml/min. The first peak corresponds to the DH/PH-RhoA complex and eluted at an apparent MW of 44 kDa. The second peak corresponds to excess RhoA and eluted at an apparent MW of 19 kDa. The gel confirms the presence of both, RhoA and the DH/PH domains in the first peak. RhoA expressed from the pMAL and the pGEX vector shows two bands on SDS-PAGE gels.

3.2.12 Crystallization of the RhoA-DH/PH complex

Crystals of the RhoA-DH/PH complex were obtained by the hanging drop vapor diffusion method in which 1 µl of protein solution was mixed with 1µl of well solution at 4 ºC. Two crystal forms, single rods and plates that were comprised of several thin densely stacked plates (see Figure 3.5) could be grown with RhoA expressed from the pGEX-KG-RhoA as well as from the pMALc2H$_{10}$T_RhoA vector. Crystals grew in a

wide range of precipitant and buffer solutions (precipitants: PEG 400, PEG 8000, PEG 10000, PEG 12000, MPEG 5000, buffers between pH 6.0 and 7.8: sodium formate, sodium phosphate, Bis Tris, MES, PIPES). Several additives were tried but did not seem to alter the formation of crystals: 1 to 10 mM EDTA, 0.5 to 2 mM $MgCl_2$, 1 to 3 % methanol, 1 % propanol, 0.4 to 1.4 M urea, 0.1 to 0.3 mM and 0.1 mM dodecyl maltoside, 5 to 15 % glycerol and 0.2 mM IP_3. Crystals could also be grown at room temperature with 15 % PEG 8000 and 10 % PEG 10000 as the precipitant. Initial crystals grew as thin short rods and optimization of the crystallization conditions resulted in thicker and longer rods. Crystals that grew as densly stacked plates could not be used for structure determination because they either cracked during harvesting or showed a highly mosaic diffraction pattern that could not be indexed. The crystal used for structure determination was grown in well solution containing 11% PEG 8K, 0.6 M NaCl, 50 mM sodium phosphate pH 7.4, and 5 mM EDTA. The presence of both RhoA and the DH/PH domains, and not just either protein in the crystals, was verified by crushing the crystals in wash buffer, washing the crystals, running the solutions on a gel and staining with silver stain. This procedure showed bands at the corresponding molecular weights for RhoA as well as the DH/PH domains. Complex crystals belong to spacegroup *C*2 and have unit cell dimensions of a = 294.9 Å, b = 95.0 Å, c = 157.0 Å and β = 94.0 º. The Matthews coefficient indicated that there are 4 to 9 complex molecules per asymmetric unit (71 to 35 % solvent content, respectively). Structure determination subsequently revealed that the complex crystallized as a tetramer with pseudo-C4 non-crystallographic symmetry and contains four molecules per asymmetric unit (V_M = 4.3 $Å^3$ Da^{-1}).

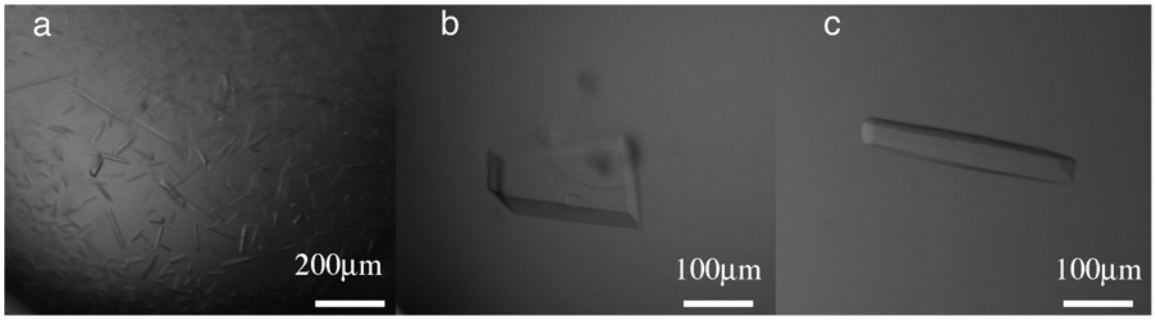

Figure 3.14: Crystals of the LARG DH/PH domains-RhoA complex. (a) Initial crystals grown as needles. (b) Initial crystals grown as stacked plates. (c) A single crystal grown under optimized conditions.

3.2.13 Harvesting of RhoA-DH/PH complex crystals

Complex crystals grew in two crystal forms that were prone to cracking during crystal harvesting. The long rods were often attached to the coverslip and cracked when cryosolution was added to the drop. The thin, stacked plates either separated during harvesting or cracked. Instead of adding cryoprotectant solution to the drop complex crystals were transferred through a series of solutions containing increasing concentrations of cryoprotectant. Several harvesting conditions had to be explored, including 20 % glycerol in combination with 10 % PEG 8000, 15 % ethylene glycol, and 25 % ribose. 15 % PEG 400 in combination with 15 % PEG 8000 seemed to retain the best diffracting crystals. The cryoprotectant solution, that was used to harvest the crystal from which the best data set was collected contained 15 % PEG 400, 15 % PEG 8000, 0.6 M NaCl, 50 mM sodium phosphate pH 7.4, 20 mM HEPES, pH 8.0, 2 mM DTT, and 5 mM EDTA. Cryoprotected crystals were flash frozen in liquid nitrogen and stored in

liquid nitrogen or propane for later data collection. Alternatively, for immediate data collection, cryoprotected crystals were flash frozen in the cryostream.

3.2.14 Data collection and processing of RhoA-DH/PH complex crystals

The initial data set of the DH/PH-RhoA complex crystals (rhoAdhph50) was collected using Cu$K\alpha$ radiation on a MAR345 imaging-plate detector (diffraction to 4.0 Å spacings). The temperature was maintained at 95 K with an Oxford Instruments Cryojet. The data set for the final structure determination (rhoAdhph92) was collected on a Quantum 210 CCD detector (ADSC) at beam line 8.2.1 of the Advanced Light Source (ALS) in Berkley, CA (diffraction to 3.2 Å spacings). The diffraction pattern was highly anisotropic with diffraction to 3.2 Å in the best direction (c axis), but only 3.75 Å and 4 Å along the a and b axes, respectively. Data were reduced and scaled using the HKL2000 package [94]. A sigma cut off of -1 was applied.

Shell Lower limit	Upper limit	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
30.00	6.87	5426.7	144.3	65.9	1.384	0.035	0.038
6.87	5.46	804.0	52.0	46.0	1.180	0.082	0.074
5.46	4.78	873.2	63.0	56.6	1.192	0.091	0.081
4.78	4.34	927.6	75.4	68.4	1.202	0.098	0.087
4.34	4.03	674.2	85.0	80.6	1.113	0.131	0.110
4.03	3.79	479.2	95.9	93.7	0.974	0.185	0.157
3.79	3.60	358.6	105.5	104.4	0.902	0.271	0.255
3.60	3.45	296.1	112.8	112.0	0.849	0.341	0.324
3.45	3.31	236.6	121.4	120.8	0.803	0.455	0.455
3.31	3.20	172.6	122.2	122.0	0.716	0.574	0.640
All reflections		1099.2	96.6	84.9	1.086	0.083	0.051

Table 3.7: Scaling statistics for the rhoAdhph92 crystals. Reflections were integrated with the program DENZO and scaled with the program SCALEPACK.

3.2.15 Structure determination of the RhoA-DH/PH complex

The 3.2 Å crystal structure of the LARG DH/PH-RhoA complex was determined by molecular replacement using the LARG DH domain (see above) modeled in complex with RhoA from the Dbs-RhoA complex structure (PDB code: 1LB1), as a search model. Molecular replacement was performed with the program MOLREP of the CCP4 suite [95]. The best solution was achieved by searching for 4 molecules per asymmetric unit. The solution had a peak height (RF/σ) of 7.03 and an initial R factor of 52 and correlation coefficient of 42. The LARG DH/PH domains in complex with RhoA could not be used as a search model, which was reflected in an R factor and correlation coefficient of 62 and 25, respectively. This is not surprising because the relative disposition of the PH and DH domain in RhoGEFs differs considerably (Fig. 3.19). The resulting electron density map allowed placement of the PH domain from the DH/PH structure (see above). Some loop regions within the DH/PH domains, for which electron density was not interpretable in the DH/PH crystal structure, showed interpretable electron density in the A and C chains of the complex crystals. This was due to stabilization of these regions by either complex formation with RhoA, in case of the N-terminal extension and the β5-β6 loop, or by crystal contacts, in case of the N-terminus of the PH domain (999-1008), the β1-β2 loop (1029-1035), the insert in the β4 strand (1062-1073), and the C-terminus of the PH domain (1136-1138) (see Fig. 3.15, 3.16, 3.18). The DH/PH domains in the complex structure were therefore used as a model to build these loop regions in the DH/PH structure. However, Asp765, residues 1002-1005 in the βN-3/10N loop, and residues

1064-1074 in the insert region in the β4 strand could not be modeled in either the complex or in the DH/PH structure.

3.2.16 Refinement of the RhoA-DH/PH complex structure

The structure was refined with the programs CNS and REFMAC with TLS refinement to final R_{work} and R_{free} of 25 and 31%, respectively (Table 3.5). Non-crystallographic symmetry (NCS) constraints were employed. Initially "tight" four-fold restraints were used and as differences between the chains became apparent, relaxed restraints were used, and ultimately removed.

Analysis of crystal packing

The DH/PH-RhoA complex crystallized as a tetramer with pseudo-C_4 non-crystallographic symmetry. The C_4 symmetric contacts between subunits involve mainly the N-terminal extension and the α4 segment and bury an unconserved tryptophan side chain Trp^{953} in the α5-α6 loop of the DH domain. This tryptophan is exposed and not ordered in the DH/PH domains structure. Although oligomers that associate with the membranes tend to have cyclic symmetry, this tetramer is not expected to be physiologically significant. Furthermore, the DH/PH-RhoA complex eluted as a monomer from gel filtration columns.

Inspection of the electron density map showed four molecules per asymmetric unit arranged as a pseudo-tetramer. The plane of the tetetramer is about 40° offset relative to the x-axis of the unit cell**.** Monomer A and C show well defined electron

density in some problem areas, especially in the loop between DH and PH domains and the N-terminal region. These two monomers are involved in crystal contacts with residues of the DH domain and RhoA. Monomers E and F do not show electron density in these areas and they are involved in crystal contacts through their PH domains. The interactions within the pseudo-tetramer are entirely through DH-DH domain contacts. The interactions between pseudo-tetramers are formed by DH-RhoA contacts as well as DH-DH and DH-PH interactions.

3.3 Structure Interpretation and Discussion

3.3.1 The LARG DH/PH domains

The core of the LARG DH domain is, like other RhoGEF DH domains, comprised of six major helical segments, in which segments 2, 3, and 5 contain kinks and are subdivided in a series of short α-helices (Fig. 3.15). The most novel structural feature of the LARG DH domain is an N-terminal extension (residues 766-781) that has not been observed in previously determined structures of DH domains. This extension forms a small subdomain, comprised of a loop and two short α-helices (αN1: residues 769-773, αN2: residues 775-781). The αN1/αN2 subdomain was highly mobile in the DH/PH structure, and was traced only after structure determination of the DH/PH-RhoA complex (see below). Sequence analysis indicates that the N-terminal extension is conserved in the Lbc family of RhoGEFs, which specifically exchange nucleotides on RhoA (Fig. 3.20). The αN1/αN2 subdomain seems to be stabilized by a hydrophobic core that centers on Trp^{769}, a residue that is invariant in the Lbc family. Furthermore, conserved residues from αN1, αN2, α1, and the α2-α3 loop participate in the hydrophobic core.

The LARG PH domain is a flattened, anti-parallel β-barrel, capped on one end by a C-terminal alpha helix (αCT) (Fig. 3.15 and 3.18). Compared with the LARG DH domain, the PH domain has substantially higher temperature factors, indicating relatively high mobility of the domain in the crystal lattice. Like other RhoGEF PH domains, the LARG PH domain has an N-terminal extension, comprised of an α-helix (αN) that is an extension of the α6 segment of the DH domain, followed by a β-strand (βN) and then a

3/10 helix ($N_{3/10}$). Residues from all secondary structure elements in the extension form a large hydrophobic core with the β2-β3 sheet, which has a high content of aliphatic hydrophobic residues. A similar hydrophobic core is observed in the Tiam, Dbs, intersectin, and SOS DHPH domains, where it is comprised of aromatic as well as aliphatic residues. The βN-strand of LARG and all other DHPH domains forms a β-sheet with the β4-strand. The PH domain of Tiam is the only RhoGEF PH domain that does not have a βN strand. However, hydrophobic residues in the α6 strand seem to be taking over the role of the βN strand and are interacting with β4 [5].

Location of the xenon and bromide sites in the DH/PH domains

To verify the validity of the determined xenon and bromide sites, their location within the DH/PH fragment was studied. Xenon atoms usually occupy hydrophobic pockets and bromide ions positively charged pockets. Two xenon sites were found in the DH domain (Xe1 and Xe3) and one xenon site in the PH domain (Xe2). The strongest peak corresponds to the xenon atom in the PH domain (Xe2). This xenon atom lies in a hydrophobic pocket made by a conserved tryptophan in the PH domain (Trp^{1123}) and two leucine residues (Leu^{1026}, Leu^{1097}). The second relatively strong peak (Xe1) corresponds to a xenon atom in the DH domain that lies in a highly conserved hydrophobic core, which holds together α1, α3, α4, and α5 and therefore stabilizes the helical bundle. The weakest peak (Xe3) corresponds to a xenon atom that lies in a hydrophobic pocket comprised of residues from the α3 and α4 helix segments (Gln^{902}, Gln^{909}, Val^{912}, Gln^{913}). The sodium bromide lies in a postively charged groove on the surface of α3, that is comprised of a proline (Pro^{892}) and a glutamine (Gln^{891}).

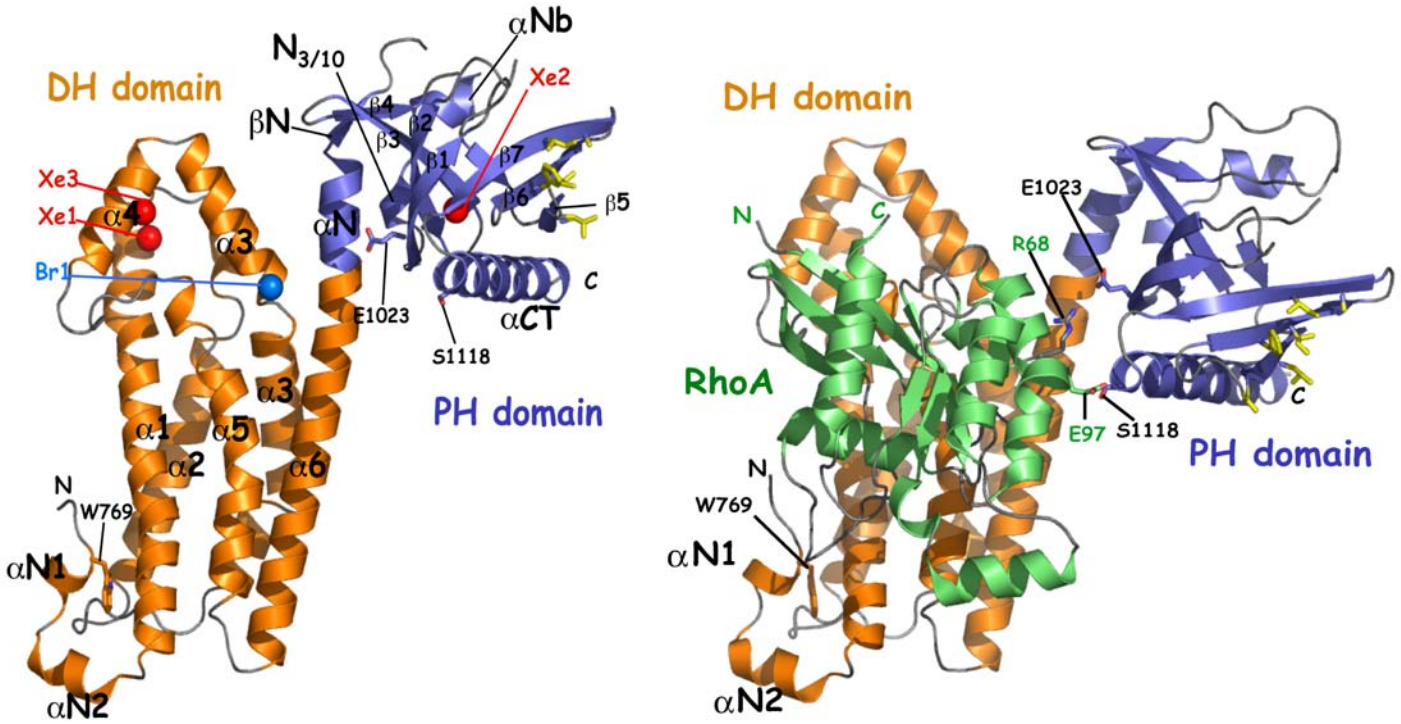

Figure 3.15: The structure of the LARG DH/PH domains (left panel) and their complex with RhoA (right panel) have been determined by MIRAS, using xenon and sodium bromide derivatives (the three xenon sites are shown as red spheres), and molecular replacement, respectively. Upon complex formation the PH domain rotates by about 30° relative to the DH domain, thus embracing RhoA. Two contact areas between the PH domain and RhoA are present that could be responsible for the PH domain-assisted nucleotide exchange seen in LARG. One is a direct contact (PH Ser^{1118}-RhoA Glu^{97}), the other one is an indirect contact where DH domain interactions with RhoA Arg^{68} are supported by a network of hydrogen bonds with the β1 strand of the PH domain. Glu^{1023} of the PH β1 strand in addition shares a long range salt bridge with Arg^{68}. An N-terminal subdomain (αN1/αN2 subdomain), that interacts with the switch I/effector loop of RhoA, has not been observed in other DH/PH domains. The subdomain is stabilized by a hydrophobic core, centered on Trp769.

3.3.2 The LARG DH/PH domains in complex with RhoA

DH/PH-RhoA crystallized as a tetramer with pseudo-C_4 non-crystallographic symmetry (Fig. 3.16). The four DH/PH domains-RhoA complexes were defined as AB, CD, EF, and GH chains. The DH/PH domains of two protomers (chains A and C) are the best ordered. The C_4 symmetric contacts between subunits bury the side chain of Trp^{953}, which occurs in the β5-β6 loop of the DH domain. Although oligomers that associate with membranes often have cyclic symmetry, this tetramer is not expected to be physiologically significant, in part because Trp^{953} is not conserved in other RH-RhoGEFs and the RhoA monomers are not consistently oriented towards the same plane.

The major contact regions with RhoA are in the DH domain. Contacts between the α1 and α5 helical segments of the DH domain and switch I of RhoA, and contacts between the α6 helical segment and switch II define the basis for the catalysis of nucleotide exchange (see 3.3.7). Contacts of the α4-α5 loop of the DH domain with the β1-β2-β3 sheet of RhoA and contacts between the α5 helical segment and the β3 strand of RhoA dictate substrate specificity (see 3.3.11).

Residues 2-181 of RhoA are observed in the most well-ordered RhoA subunit (B) of the tetramer, and the backbone conformation of RhoA is essentially the same as that of RhoA in complex with Dbs (0.6 Å r.m.s. deviation) [3]. The changes seen in DH/PH-bound RhoA compared to nucleotide-bound RhoA are in the switch I and switch II of RhoA, as well as the P-loop and the ^{116}GXKXDL^{121} motif. There are two structures of GDP bound RhoA, one including Mg^{2+} [106] and one without Mg^{2+} [1]. Comparison of

these structures reveals expected conformational changes in the switch II region, which plays an important role in coordinating the Mg^{2+} ion. In the RhoA-GDP structure (without Mg^{2+}), switch II is collapsed into the Mg^{2+} and γ-phosphate binding sites. A similar conformation of switch II is seen in the complex structure with the LARG DH/PH domains. The conformation of switch I in the RhoA-GDP structure was unexpectedly different from the RhoA-GDP-Mg^{2+} structure. Switch I is in a position similar to what has been observed for switch I of Ras in the Ras-SOS structure [2], where it is completely removed from the nucleotide binding site (the largest displacement is 19 Å for Val^{33}). Switch I in the DH/PH-RhoA structure is in a similar orientation as in the RhoA-GDP-Mg^{2+} structure, although interactions of switch I with the LARG αN1/αN2 subdomain leave the nucleotide binding pocket in a slightly more exposed position compared to RhoA-GDP-Mg^{2+} (displacement for Val^{33} is 6.8 Å).

Switch I has increased B-factors compared to the remaining residues of RhoA, which indicates high mobility of this region. Increased B-factors in switch I have been observed in other GTPases in complex with DH/PH domains. However, in these structures the increase of B-factors in this region is much more pronounced, indicating that the N-terminal extension in the LARG DH/PH domains might contribute to the stability of this region. The insertion region of RhoA (Fig. 1.1), which is implicated in effector binding [33, 34] is completely exposed to solvent and indicates that the formation of a signaling complex is possible, in which RhoA can bind effector molecules while it is activated by the DH/PH domains. The nucleotide binding site of RhoA showed strong positive electron density in the $2F_o$-F_c map in place where the α-phosphate

is located in nucleotide bound RhoA. The density was therefore modeled as a phosphate because crystals were obtained in phosphate buffer (3.2.5).

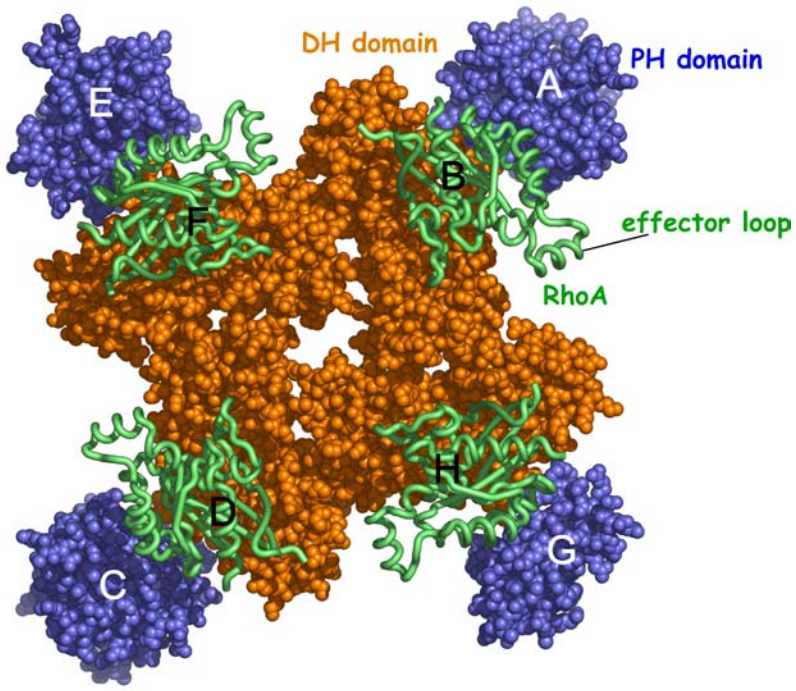

Figure 3.16: The DH/PH-RhoA complex crystallized as a tetramer with pseudo-C_4 non-crystallographic symmetry. The four DH/PH domains-RhoA complexes were defined as AB, CD, EF, and GH. The DH/PH domains of two protomers (A and C) are relatively well ordered compared with those of the DH/PH structure. This tetramer is not expected to be physiologically significant, because the symmetry contacts between subunits are not conserved in other RH-RhoGEFs, the RhoA monomers are not consistently oriented towards the same plane, and the complex elutes as a monomer from a gel filtration column. The insertion region of RhoA, which is implicated in effector binding [33, 34] is completely exposed to solvent and indicates that the formation of a signaling complex is possible, in which RhoA can bind effector molecules while it is activated by the DH/PH domains.

3.3.3 Changes in the LARG DH/PH domains upon complex formation

The overall conformation of each of the individual DH and PH domains in the complex is similar to their counterparts in the DH/PH structure (r.m.s. deviations of 0.995 and 0.749, respectively). Conformational changes in the DH domain occur in the N-terminal extension, the α4-α5 loop, and the DH/PH junction. These regions directly engage RhoA. The N-terminal extension is pulled towards RhoA and residues Asn^{768} and Glu^{790} are re-oriented to provide a more compatible surface for the interaction with RhoA (Fig. 3.17). The α4-α5 loop is pushed away from RhoA by about 1.6 Å due to steric interference with the RhoA N-terminus and the RhoA β2-β3 sheet.

Side chains of the DH domain in the DH-RhoA binding interface are essentially unchanged, except for two residues in the α5 and α6 segments (Gln^{935}, Asn^{975}, respectively), whose side chains are altered to provide a more optimal binding interface for the RhoA residues Asn^{41} and Arg^{68}, in the effector loop and switch II, respectively. Both, Gln^{935} and Asn^{975} are highly conserved in the Dbl family DH domains and mutations of these residues in the Dbl and Trio DH domains severely impaired nucleotide exchange activity [8, 9]. The conformation of the PH domain is essentially unchanged, although the loops have become better ordered.

Changes in the DH-PH junction

The structures of the LARG DH/PH domains in the unbound form as well as bound to their substrate GTPase RhoA have been determined. The two structures now

provide evidence for the proposed importance of the C-terminal region of the Tiam DH domain in nucleotide exchange [5] and for the severely diminished nucleotide exchange activity of mutants in the DH-PH junction of Dbl [107]. Although the conformation of the individual DH and PH domains in the complex structure are similar to their equivalents in the DH/PH structure, the DH-PH junction shows striking differences. Upon complex formation with RhoA the PH domain is collapsed relative to the DH domain, by approximately 30˚, as if the two domains are embracing RhoA (Figure 3.15). This conformational change is similar in all four complexes of the crystal structure and mainly involves the hinge between the DH and PH domains at the end of the α6 helix spanning residues 975 to 986. Because each complex has a unique crystalline environment, the observed collapse is not likely an artifact of crystal packing. Each of the PH domains adopts a slightly different orientation with respect to the DH domain. However, each PH domain maintains similar contacts with RhoA. Two scenarios can be envisioned that would cause the observed conformational change of the DH/PH domains upon binding RhoA. Firstly, interactions of Arg^{68} in RhoA with Glu^{982}, which is highly conserved throughout the Dbl family, and with Asn^{983}, could cause the PH domain to be rotated (Fig. 3.17). Secondly, the PH domain could rotate upon binding to either membranes or other regulatory proteins, which then could favor RhoA binding. The rotation of the PH domain causes one new contact to be established between the DH and the PH domain that is not present in the DH/PH structure. Lys^{979} hydrogen bonds to the backbone carbonyl of Ile^{1021} in the β1 strand of the PH domain. His^{814} in Dbs is the equivalent residue to Lys^{979} in LARG and is also involved in contacts with the Dbs PH

domain (Tyr^{889}) [4]. His^{814} in Dbs forms additional contacts with RhoA and mutating either His^{814} or Tyr^{889} resulted in highly diminished exchange activities, confirming the importance of DH-PH domain contacts in nucleotide exchange.

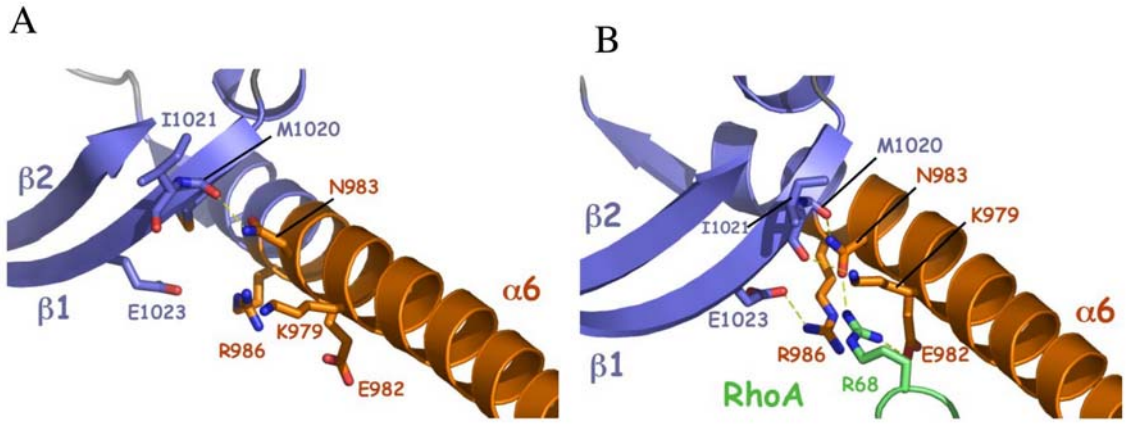

Figure 3.17: Changes at the DH-PH junction upon complex formation. Upon complex formation with RhoA, the PH domain collapses relative to the DH domain and forms indirect (via the DH domain) contacts with RhoA at the DH-PH junction. Residues of the β1 strand of the PH domain interact with C-terminal residues of the α6 segment that directly contact Arg^{68} of RhoA (PH Met^{1020}-DH Asn^{983}, PH Glu^{1023}-DH Arg^{986}). A similar role of the PH domain in supporting interactions of the DH domain with RhoA has been seen in Dbs and confirmed by site directed mutagenesis [4]. The rotation of the PH domain causes one new contact to be established between the DH and the PH domain that is not present in the DH/PH structure. Lys^{979} hydrogen bonds to the backbone carbonyl of Ile^{1021} in the β1 strand of the PH domain. The equivalent residue in Dbs, which also exhibits PH domain-assisted nucleotide exchange, interacts with a PH domain residue and is crucial for its catalytic activity [4].

.

3.3.4 The LARG DH and PH domains in comparison to other RhoGEFs

The DH domain

The most distinct structural feature of the LARG DH domain is an N-terminal subdomain that might be present throughout the Lbc family RhoGEFs (Fig. 3.10). Intersectin is an outlier in this family, which specifically exchanges nucleotides on Cdc42. Although Trp^{769}, a residue at the center of the hydrophobic core of the subdomain, is also conserved in intersectin, intersectin has a two amino acid deletion in the N-terminal extension (residues equivalent to LARG Leu^{772} and Val^{773}) and several substitutions (arginine substituted for Ile^{786}, glycine substituted for Glu^{790}) that could abrogate the formation of an equivalent N-terminal subdomain. The construct that was used in the structure determination of the intersectin DH/PH domains did not contain the entire N-terminal extension [3]. However, the three residues N-terminal to α1 in intersectin are superimposable onto αN_1 of LARG.

When superimposed onto the Dbs [3, 4], intersectin [3], and Tiam [5] DH domains the most prominent structural variations in the LARG DH domain are in the α2-α3 and the α4-α5 loop, as well as in the C-terminus of α6 (Fig. 3.16). The α2-α3 loop in the LARG and intersectin DH domains has a six residue insertion with respect to the Dbs DH domain. In LARG this extension allows the α2-α3 loop to contribute a hydrophobic residue (Val^{860}) to the N-terminal subdomain (Fig. 3.10). In intersectin, Val^{860} is substituted with a proline, which could also contribute to a hydrophobic core and stabilize a subdomain. The α2-α3 loop of Tiam has a 15 residue insertion compared to the Dbs

DH domain, which forms a short 3/10 helix. The functional implications of this extended α2-α3 loop in Tiam are not obvious from the structure.

The α4-α5 loop is one of the primary specificity determining regions of the DH domain (see 3.3.8) and accommodates the N-terminal region of the GTPase including Arg^5, which is substituted by shorter residues in Rac1 (Ala^3) and Cdc42 (Thr^3). In the LARG structure the segment encompassing the α4-α5 loop, the α4 helix and the C-terminal part of the α3 helix, is further away from RhoA by about 2.5 Å compared to the Tiam-Rac1 and Dbs-Cdc42 structures. The N-terminus of the GTPase therefore seems to determine the position of this region relative to the helical bundle.

The C-terminus of the α6 segments of RhoGEF DH domains constitutes the hinge region between the DH and the adjacent PH domain and the structural variations in this region determine the role of the PH domain in nucleotide exchange (see 3.3.3 and 3.3.9). The α6 segments of Dbs and Tiam are about 2.5 turns shorter than that of LARG and its closest homolog intersectin. This allows the adjacent PH domain in Dbs to make direct contacts with its substrate GTPase. These contacts are neccessary for full catalytic activity of Dbs *in vivo* [88] (see 3.3.9). However, such contacts were not observed in the Tiam-Rac1 structure, due to a completely different arrangement of the Tiam PH domain relative to its DH domain. Although, the α6 segment of the LARG DH domain is longer than the one in Dbs, a direct contact of the PH domain with RhoA is present due to a 7 Å translation of the PH domain towards RhoA compared to the Dbs PH domain (see 3.3.9).

The LARG DH domain has a three-residue insertion (Gly^{874}, Ala^{875}, Gly^{876}) that forms a bulge at the junction between α3a and α3b. This insertion is present throughout the Lbc family, with the exception of intersectin. However, the residues that comprise this insertion are not conserved.

The PH domain

The PH domain of LARG has notable structural differences from other RhoGEF PH domains. The β3-β4 loop of LARG is an abrupt β-turn, whereas in the Dbs-Cdc42 and Tiam-Rac1 complexes the analogous loop is extended by 7 and 25 amino acids, respectively. In the case of Dbs, the loop forms specific contacts with Cdc42 (Fig. 3.19). While Dbs has a continuous β4 strand, intersectin has a three-residue bulge (residues 1515 to 1517) and LARG has a disordered 18-residue loop (residues 1060 to 1077) (Fig. 3.18). The function of this insertion in LARG and other RH-RhoGEFs is not known, although it projects towards what is anticipated to be the membrane binding surface of the PH domain. It therefore could be involved in membrane association. Among RhoGEF PH domains, the loop connecting βN and $N_{3/10}$ is highly variable. In the case of LARG, it consists of an 18-residue loop that contains an extra helix (αNb, residues 1005-1013).

The region between the end of αNb and the β1 strand as well as most of the β1 strand show a high degree of sequence conservation in the family of RH-RhoGEFs. This region is involved in contacts with the α6 segment and contains many residues that comprise the hinge region between the DH and the PH domain (Fig. 3.17). This region

therefore determines the relative arrangement of PH to DH domains, which might be of importance for optimal orientation of the RhoGEF at the plasma membrane with respect to RhoA.

Little is known about the membrane association and phospholipid binding capacities of RH-RhoGEF PH domains. Results from a modeling study by Blomberg and Nilges [108] suggested that DH domain-associated PH domains (*i.e.* Lbc, Cdc24, Ect2, Unc-89, Scd1) might have a reverse electrostatic surface potential compared to PH domains that are known to bind phospholipids. Modeling of these RhoGEF PH domains predicted that the regions, usually involved in phospholipid binding, would have a strong negative instead of a positive potential. This prediction was confirmed for the Unc-89 PH domain, whose structure revealed that one of the known phosphoinositol binding sites between the β1-β2 and β3-β4 loops is occluded and lined by hydrophobic and negatively charged residues [109]. This study suggests that RhoGEF PH domains mainly serve as protein-protein interaction domains rather than membrane association domains. However, the LARG PH domain exhibits an electrostatic surface potential similar to PH domains that are known to bind phospholipids. The β1-β2 loop and the β5 strand have a positive surface potential. Also the region at the base of the β4 insertion is positively charged (Figure 3.18). Although, LARG and other RhoGEF PH domains do not posses a high-affinity 3-phosphoinosite binding motif that was proposed previously [110], the intersectin, Dbs and Tiam PH domains were found to bind phospholipids with low affinity (5-10 μM) [111]. The LARG PH domain has a possible binding site for $IP(1,4,5)_3$ (Fig. 3.18) that is located between the β1-β2 loop and the β3-β4 loop and

comprised of β4-strand resdidues (Gln^{1051}, Gln^{1053}, Arg^{1056}, Arg^{1060}) that are highly conserved in the Lbc family of RhoGEFs.

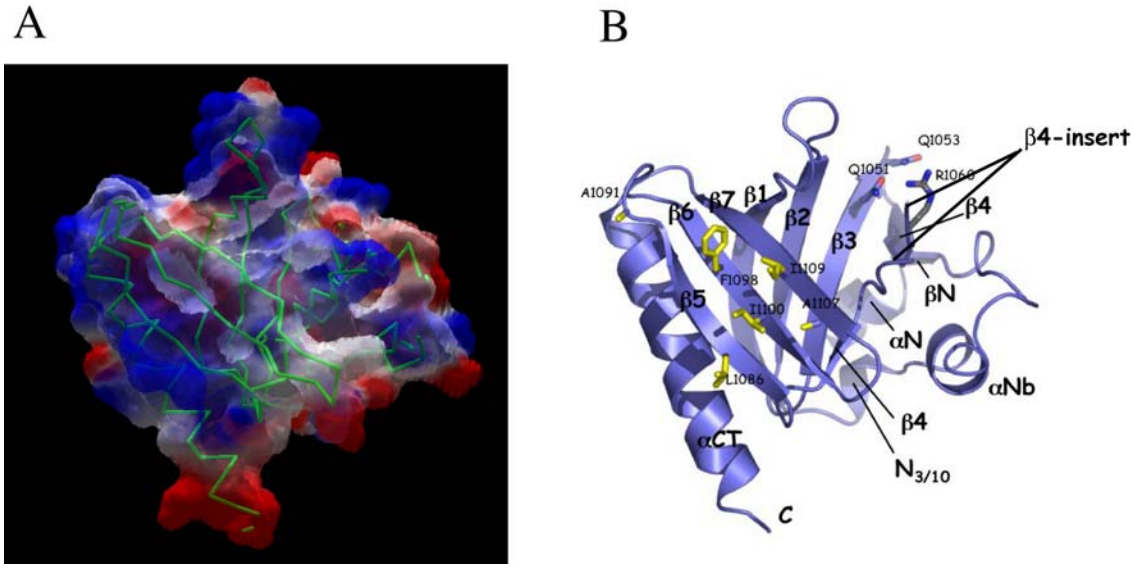

Figure 3. 18: (A) Electrostatic surface potential of the LARG PH domain. The electrostatic surface potential was generated with Swiss pdb viewer [112] and is contoured at ±1.5 kT. The alpha carbon trace of the PH domain is shown in green. The electrostatic surface potential of the LARG PH domain is similar to the one of PH domains that are known to interact with membranes. **(B) The structural features of the LARG PH domain.** Like other RhoGEF PH domains the LARG PH domain has an N-terminal extension comprised of an α-helix (αN) that is an extension of the α6 segment of the DH domain, followed by a β-strand (βN) and then a 3/10 helix ($N_{3/10}$). The LARG PH domain has a conserved positively charged patch (Gln^{1051}, Gln^{1053}, Arg^{1060}) that could interact with inositol phosphates, and a conserved hydrophobic patch (yellow side chains, Leu^{1086}, Ala^{1091}, Phe^{1098}, Ile^{1100}, Ala^{1107}, Ile^{1109}) that could be used as a docking site for regulatory proteins. This hydrophobic patch is involved in crystal contacts in both the DH/PH and the DH/PH-RhoA complex crystals.

The relative orientation of PH to DH domains

The relative orientation of PH to DH domains in RhoGEFs varies, and seems to dictate the involvement of the PH domain in catalysis. PH domain-assisted nucleotide exchange *in vitro* has been shown for Dbs [4] and LARG [54]. In both structures the PH domains are in similar orientations and have direct and indirect (via the DH domain) contacts with the GTPase. This is in contrast to the intersectin PH domain, which does not contact the GTPase and does not assist nucleotide exchange *in vitro* [3, 113]. The α6 segment of LARG is by 10 residues longer than the one in Dbs, or about 2.5 helical turns, and thus the PH domain of LARG is translated about 13 Å along the length of the α6 segment compared to that of Dbs (Fig. 3.16). However, the LARG PH domain is also translated by about 7 Å towards the GTPase, relative to the Dbs PH domain. This allows the LARG PH domain to make direct interactions with RhoA via the N-terminus of its αCT. The PH domain of intersectin is similarily translated along the α6 segment as that of LARG, however, it is oriented differently and does not have direct contacts with Cdc42, neither does it support contacts of the DH domain with Cdc42 [3]. While *in vitro* studies confirm that the intersectin PH domain does not play a role in Cdc42 binding or catalysis, *in vivo* the PH domain is necessary for full catalytic activity [113]. This suggests that regulation of the intersectin PH domain, possibly by membrane binding or other signaling domains or proteins, contributes to DH/PH activity *in vivo*, which might be a mechanism that is generally used by RhoGEFs.

The Tiam PH domain in the Tiam-Rac1 complex is rotated away from the GTPase so that its $N_{3/10}$ helix makes contacts with residues from the α3 and α6 segments

of the DH domain [5]. These contacts are conserved in the Sif GEFs, which are close homologs to Tiam and hints against the possibility that they are caused by crystal packing artifacts. The most divergent arrangement between PH and DH domains is seen in the SOS structure, where the PH domain is rotated towards the α1 and α5 segments of the DH domain and occludes the GTPase binding site [6]. Because the SOS DH/PH fragment is not active *in vitro* and only activated *in vivo* when co-expressed with activated Ras [86], the structure could represent a DH domain that is auto-inhibited by its adjacent PH domain. It is believed that the auto-inhibition is relieved upon binding of PI(3, 4, 5)P_3, a product of Ras-mediated activation of phosphoinosite 3-kinase [86].

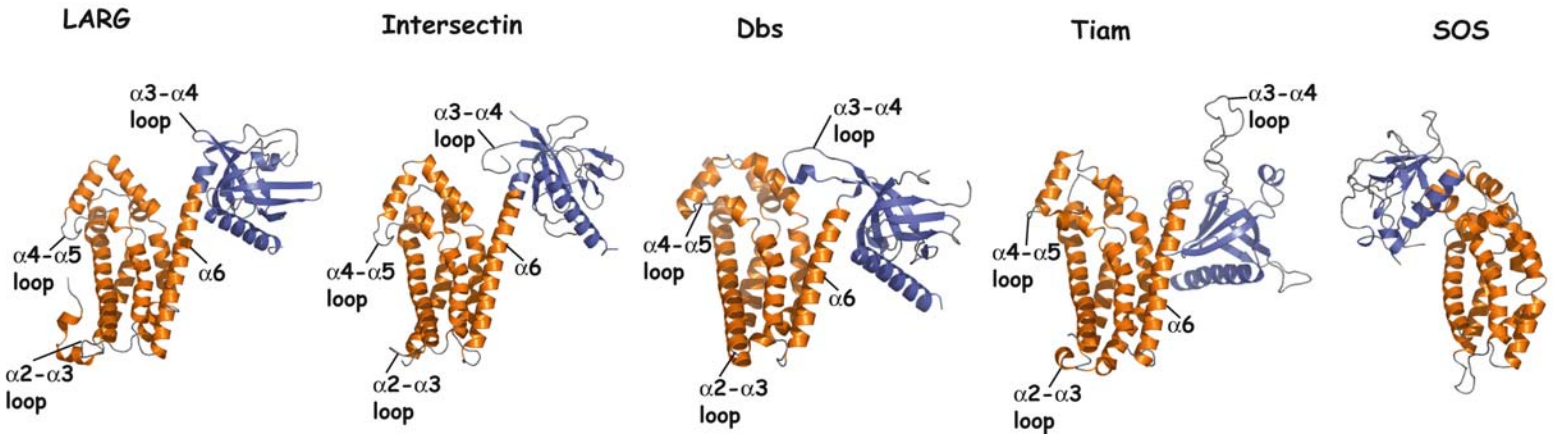

Figure 3.19: The relative arrangements of PH and DH domains in RhoGEFs in complex with their substrate GTPases (not shown). DH domains are colored orange, PH domains are colored blue. Although intersectin has the highest sequence homology to LARG, the relative arrangement of the PH to the DH domain in LARG is closest to the Dbs DH/PH domains. Both, LARG and Dbs exhibit PH domain-assisted nucleotide exchange *in vitro* and their PH domains directly contact the substrate GTPase. The intersectin PH domain does not have contacts with Cdc42 and PH domain-assisted nucleotide exchange was not detected *in vitro*. The Tiam PH domain is rotated towards the back side of the helical bundle away from Rac1 and information on the contribution of its PH domain to nucleotide exchange is not available. The SOS DH/PH domains are the only example so far for an auto-inhibitory role of a PH domain on its adjacent DH domain. PDB accession codes are LARG: 1TXD, Intersectin: 1KI1, Dbs:1LB1 and 1KZ7, Tiam: 1FOE,

3.3.5 The αN1/αN2 subdomain

Sequence analysis indicates that the αN1/αN2 subdomain is conserved in the Lbc family of RhoGEFs, which specifically exchange nucleotides on RhoA (Fig. 3.10). The αN1/αN2 subdomain seems to be stabilized by a hydrophobic core that centers on Trp^{769}, a residue that is highly conserved in the Lbc family (Fig. 3.17). Furthermore, residues from αN_1 (Val^{773}), αN_2 (Val^{777}, Leu^{778}, Leu^{781}), α1 (Ile^{786}, Gln^{789}, Glu^{790}), and the α2-α3 loop (Ser^{859}, Val^{860}) participate in stabilizing the hydrophobic core.

The αN1/αN2 subdomain interacts with the effector loop of RhoA (Lys^{27}, Val^{33})

Sequence analysis indicates that the αN1/αN2 subdomain of LARG is highly conserved among the Lbc subfamily of RhoGEFs, nearly all of which are selective for RhoA (Fig. 3.17), indicating that it might be specificity determining. Interestingly, this subdomain interacts with the effector loop/switch I of RhoA and this interaction might be the basis for the more pronounced exposure of the nucleotide binding pocket in the LARG DH/PH domains compared to other DH/PH domains that do not posses an N-terminal extension. This could indicate that the subdomain is necessary for catalysis by promoting the formation of the nucleotide-free form of RhoA. The principal contacts are between Val^{33} of RhoA, a residue conserved as glutamate in other Rho GTPases, and the side chains of Asn^{768} and Gln^{770} of αN1, and between Tyr^{34} and Glu^{790} of the α1 helix (Fig. 3.17). Although disordered in the crystal structure, the side chain of Lys^{27} is in close proximity to the backbone carbonyls of Pro^{766} and Pro^{767} of LARG. Lys^{27} is

conserved as threonine in other GTPases. Therefore, LARG potentially has an additional GTPase-selecting determinant at the N-terminus of its DH domain. In support of this hypothesis, an interesting observation was made when an LARG DH/PH domain construct (residues 785-1140), that lacked the N-terminal extension, was studied [54]. This construct showed some nucleotide exchange on Cdc42. If the N-terminus was indeed specificity determining, it would suggest that intersectin, which is the only member of the Lbc family that is specific for Cdc42 and not RhoA, does not have this N-terminal subdomain. Intersectin has a two-residue deletion within the N-terminal extension with respect to LARG, and several amino acid substitutions in α1 (a glycine substituted for Glu^{790} and arginine substituted for Ile^{786} in LARG) that could potentially abrogate formation of an analogous N-terminal subdomain. In the crystal structure of the intersectin-Cdc42 complex, only some of the residues corresponding to the N-terminal extension (residues 1229 to 1232) were present in the recombinant protein used for structure determination. Even so, these four residues are superimposable with their equivalents in the LARG N-terminal subdomain.

A residue of the αN1/αN2 subdomain (Glu^{790}) exhibits an altered conformation upon complex formation

Compared to the unbound DH/PH domains, Glu^{790} dramatically alters its conformation upon complex formation with RhoA. Glu^{790} participates in a hydrogen bonding network with the backbone amides of Trp^{769} and Gln^{770} as well as with the

Gln^{770} side chain and packs against Tyr^{34} of switch I, which is sandwiched between Glu^{790} and Pro^{32} of RhoA. The changed conformation of Glu^{790} seems to be a result of steric interference with the backbone carbonyl of Val^{33} of RhoA. The altered conformation of Glu^{790} might be structurally necessary for two reasons. First, the interactions of Glu^{790} with the N-terminal subdomain might be necessary for its stabilization. Indeed, αN1 has significantly lower B-factors in the complex structure. Second, the N-terminal subdomain might be necessary to restrict the conformation of Glu^{790}, to a compatible conformation for RhoA binding. The lack of an N-terminal subdomain might leave Glu^{790} in an extended conformation that is not compatible with RhoA binding or does not provide a beneficial binding surface for Tyr^{34}. Interestingly, Glu^{790} is conserved in the Lbc family except for intersectin, where is a glycine (Gly^{1240}). This substitution leads to a slightly different conformation of Tyr^{32} in Cdc42 (equivalent to Tyr^{34} in RhoA), which is now sandwiched between the backbone of Gly^{1240} and Pro^{30} (equivalent to Pro^{32} in RhoA). The van der Waals interaction of Glu^{790} with Val^{33} might be specificity determining because Val^{33} is substituted by a glutamic acid in Cdc42, which would cause steric repulsion. The substitution of this residue for a glycine in intersectin might be necessary for its interaction with Cdc42.

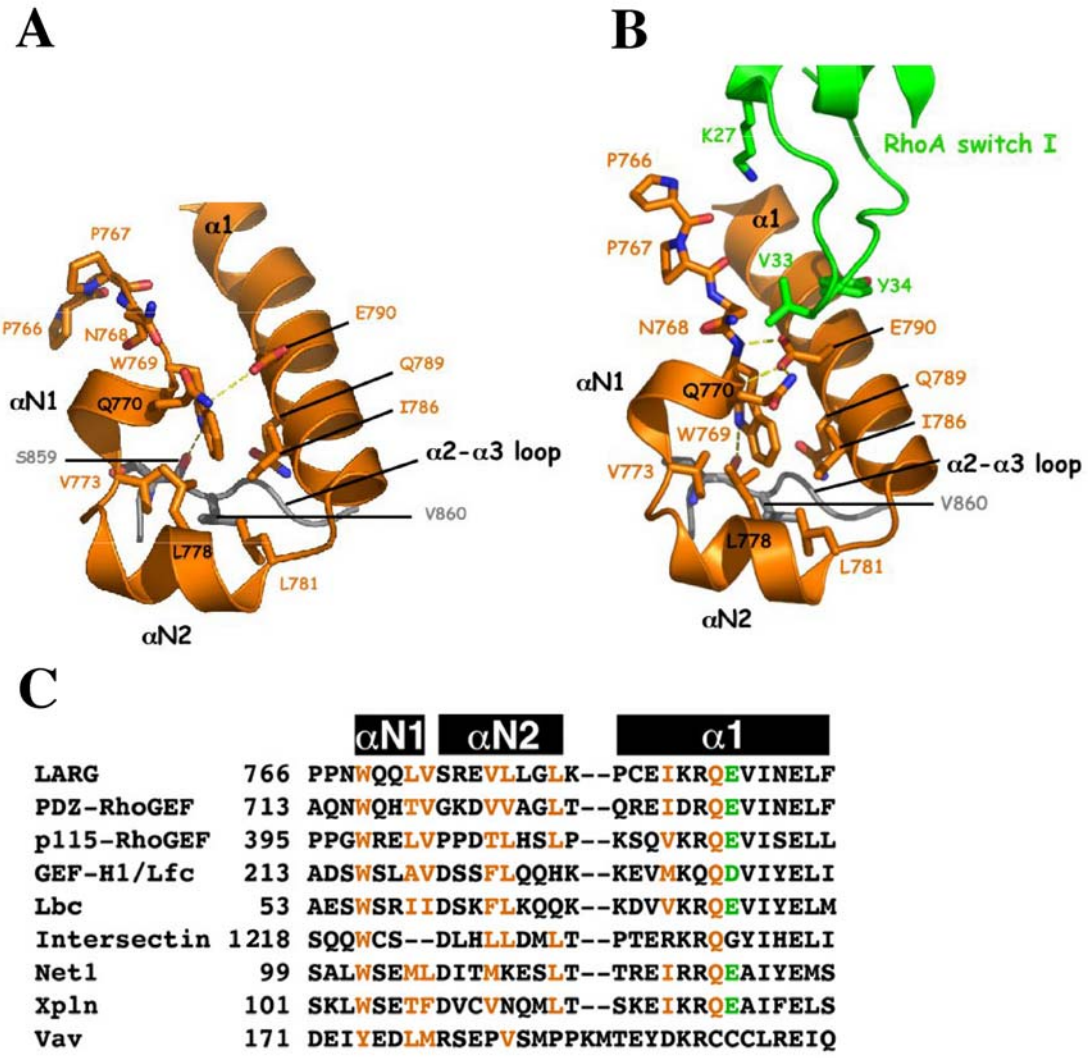

Figure 3.20: (A) A novel subdomain, the αN1/αN2 subdomain, has been revealed at the N-terminus of the LARG DH/PH domains. The αN1/αN2 subdomain contains a hydrophobic core, that is comprised of residues from the αN1 and αN2 helices as well as the α2-α3 loop, and is centered on Trp^{769}. These residues are conserved in the Lbc family of RhoA specific RhoGEFs. Intersectin is an outlier in this family because it is specific for Cdc42 and it has a two residue deletion and several substitutions that could abrogate the formation of an equivalent subdomain. **(B) The αN1/αN2 subdomain interacts with switch I of RhoA.** Glu^{790} packs agains Val^{33} and Tyr^{34} and changes its conformation upon complex formation. Glu^{790} is substituted with a glycine in intersectin, which might be necessary for its specificity for Cdc42. **(C) Sequence alignment of the region N-terminal to the DH domain in Lbc family RhoGEFs and Vav.**

3.3.6 Structural basis for LARG catalyzed nucleotide exchange

Like other GEFs, LARG does not directly contribute residues that help to eject the the nucleotide, instead it changes the magnesium and nucleotide coordinating sites of the GTPase to favor nucleotide release and stabilize the nucleotide free form of the GTPase. The LARG-RhoA interface is made primarily by contacts between residues in the α1, α5, and α6 segments of the DH domain with switch I and II of RhoA, as observed in previous GEF-GTPase structure. These contacts are highly conserved and define the basis for disruption of the magnesium binding site and exposure of the nucleotide binding site of RhoA. Upon complex formation, switch I and switch II of RhoA are reorganized such that nucleotide release is favored (Fig. 3.18). Switch II is reorganized mainly by hydrophobic contacts between Tyr^{66} in RhoA and the α5 and α6 segments. The side chains of Ala^{61} and Glu^{64} in switch II are re-oriented into the magnesium and γ-phosphate binding sites, respectively (Fig. 3.18). Glu^{64} is now in hydrogen bonding distance to Lys^{18} of the P-loop. Lys^{18} and Glu^{64} are highly conserved in small GTPases and are crucial for nucleotide exchange reactions [114, 115] possibly due to their contribution to stabilizing the nucleotide-free state of the GTPase. Switch I, which in the GDP bound form of RhoA coordinates the magnesium and nucleotide binding sites, is re-arranged by interactions with the α1, α5a and αN2 helices of the DH domain, resulting in exposure of the nucleotide binding site. Glu^{935}, Lys^{939}, and Leu^{943}, three highly conserved residues in the α5 segment of DH domains, would sterically interfere with the switch I loop in the GDP bound form of RhoA. It therefore, appears as if these residues push switch I away from the DH domain. The changed conformation of switch I seems to be stabilized by

interactions with α5a1 and αN2, involving interactions of Glu^{794} of α1 with the backbone amides of Thr^{37} and Val^{38} of RhoA. This glutamic acid is highly conserved in RhoGEFs, and is also the major contact residue between the DH/PH domains and switch I in the Tiam-Rac1 [5], Dbs-RhoA [3], and intersectin-Cdc42 structures [3]. Thr^{37} of switch I, which coordinates the magnesium ion in the Rho-GDP-Mg^{2+} structure [106] (Fig. 3.18) is in hygrogen bonding distance to Asn^{946} and removed from the magnesium binding site. The interactions of the N-terminal subdomain in LARG with switch I are probably responsible for a slightly more exposed nucleotide binding site of RhoA compared to the other DH/PH-GTPase complex structures, where no N-terminal extension was observed.

In previously determined DH/PH-GTPase structures the only structural basis for nucleotide expulsion is interference with the magnesium binding site by switch II residues and exposure of the nucleotide binding site due to re-arrangement of switch I (Fig. 3.18). The nucleotide binding pocket in these structures is otherwise undisturbed. However, in the LARG DH/PH-RhoA complex there are also interferences with the phosphate and purine binding sites of the nucleotide binding pocket. The backbone carbonyl of Gly^{14} in the P-loop (phosphate binding loop) sterically and electrostatically occludes the β-phosphate binding site. The purine binding site of RhoA is altered by re-orientation of Asn^{117} towards the purine. However, the functional implication of the re-arrangment of Asn^{117} is not clear because it seems to be favourably oriented to participate in a hydrogen bond with N7 of the purine. A similar conformation of the residue equivalent to is observed in the Ras-SOS [2] and EF-Tu•EF-Ts [14] complex structure.

The collapse of Gly^{14} into the phosphate binding pocket is notable, because a similar collapse of the P-loop has only been observed in complexes with GEFs that do not contain a DH/PH module and directly contact the P-loop (*i.e.* EF-Ts-EF-Tu, SOS-Ras, Sec7-Arf). In contrast, Tiam, Dbs, and intersectin do not interact with the P-loop and in the structures of the DH/PH domains of these RhoGEFs with their respective GTPases the P-loop is intact and the β-phosphate binding site is coordinated with sulfate ions or highly ordered water molecules. The intact P-loop in the RhoGEF structures could explain the about one to two orders of magnitude lower catalytic activities of RhoGEFs compared to other GEFs. However, in the recent structure of Ran in complex with its GEF RCC1, the P-loop of Ran is also intact and stabilized by sulfate ions, even though there are direct contacts of the GEF with the P-loop. This suggests that GEF catalyzed reactions generally proceed through the same structural changes independent of interactions of the GEF with the P-loop of the GTPase, and that the crystallization conditions might dictate if the GTPase is trapped with an intact or collapsed P-loop. This is confirmed by the LARG-RhoA complex structure, where the P-loop is collapsed although there are no direct contacts between the GEF and the P-loop. It was suggested that structures with an intact P-loop mimic a low-affinity ternary complex intermediate of GEF, GTPase and loosley bound nucleotide, whereas structures with a collapsed P-loop represent the high-affinity binary complex of GEF and GTPase [10, 11].

The structural basis for the re-orientation of Gly^{14} and Asn^{117} in the complex with the LARG DH/PH domains is not clear, but it might be a result of the interaction of the N-terminal subdomain with switch I. This interaction could stabilize the exposed

conformation of the nucleotide binding pocket. Compared to other DH/PH-GTPase complexes this might not allow buffer ions to be trapped in the phosphate binding sites and the P-loop might therefore collapse into a conformation similar to the one seen in the LARG-RhoA complex.

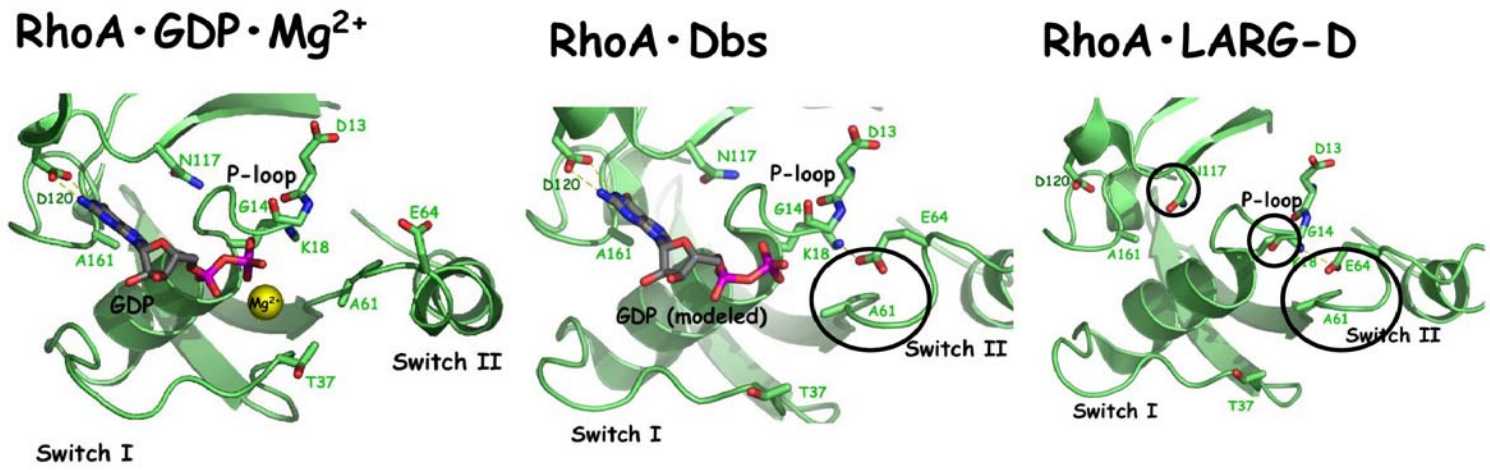

Figure 3.21: Changes in the nucleotide binding site upon complex formation with DH/PH domains. The coordination of nucleotide and magnesium in the RhoA•GDP•Mg^{2+} structure is shown in the first panel. In all previously determined DH/PH•RhoA complex structures only interference with the magnesium binding site (Ala^{61}) and with the γ-phosphate binding pocket (Glu^{64}) was observed, resulting in a conformation that is still compatible with GDP binding. As an example RhoA in the DbsRhoA complex structure is shown in the second panel. GDP was modeled from the magnesium-free RhoA structure [1]. In contrast, the nucleotide binding pocket of RhoA in complex with the LARG DH/PH domains (third panel) is not compatible with GDP binding because the carbonyl of Gly^{14} is re-oriented in to the β-phosphate binding site and Asn^{117} points into the purine binding pocket.

3.3.7 An alternative conformation of RhoA

There are four unique DH/PH-RhoA complexes in the asymmetric unit of the DH/PH-RhoA crystals (complexes AB, CD, EF, GH). In one of the complexes (AB), crystal contacts between RhoA-Asp^{165} and the α3 segment of a symmetry related AB complex have trapped RhoA in a dramatically different conformation. Residues 160-164 of RhoA, particularly Ala^{161}, have collapsed into the purine-binding pocket (Fig. 3.22). An analogous collapse is observed in the Ras-Sos complex [2]. The side chain of RhoA-Asp^{120}, which in nucleotide-bound structures of RhoA forms hydrogen bonds with the N1 and N2 nitrogens of guanine, is unchanged but instead supports the new conformation of this loop with two backbone hydrogen bonds. While this conformation is caused by crystal contacts, it might represent an intermediary state of RhoA, after loss of magnesium and nucleotide. It can be envisioned that this complex then relaxes into a conformation similar to those of the three other complexes in the asymmetric unit, where the nucleotide binding pocket is open and only the phosphate binding site is occluded. The nucleotide would then bind with the purine first and subsequently establish the contacts with the β- and γ-phosphates and ultimately lead to dissociation of the DH/PH domains from RhoA. A similar sequence of nucleotide re-entry has been proposed for the SOS catalyzed nucleotide exchange on Ras [2]. However, for the RCC1 catalyzed exchange on Ran it has been proposed that the phosphate binds first and then the nucleotide based on the fact that the P-loop is intact [11].

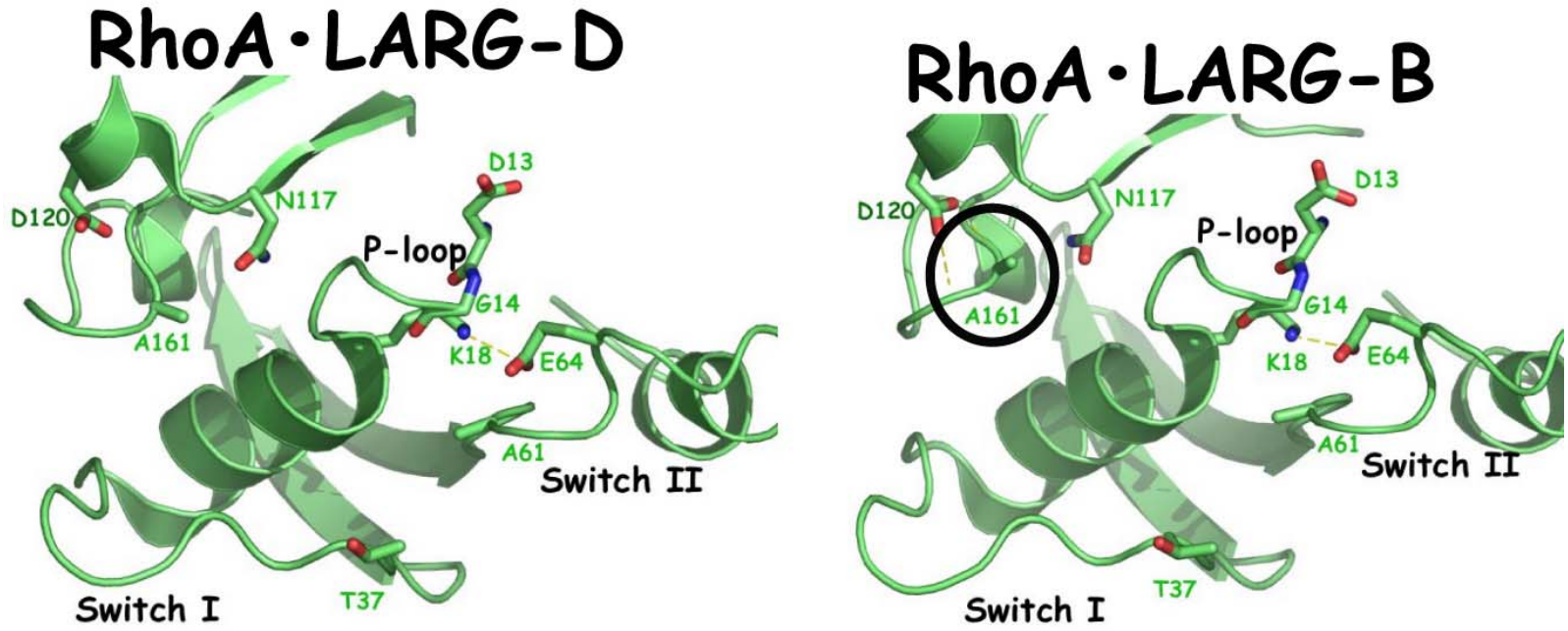

Figure 3.22: Two different conformations of the nucleotide binding pocket of RhoA in complex with the LARG DH/PH domains were observed. Crystal contacts left the nucleotide binding pocket of RhoA in one of the complexes (AB) in the tetramer in an alternative conformation compared to the other three complexes (CD, EF, GH). Ala^{161} of the SAK motif is pointed into the purine binding site, further adding to the incompatibility of RhoA with GDP binding. This conformation is supported by Asp^{120} of the GXKXDL motif, which in nucleotide bound RhoA coordinates the base. Although a result of crystal contacts this conformation could be a possible conformation RhoA assumes during the expulsion of nucleotide and it reflects the plasticity of RhoA when it is bound to DH/PH domains.

3.3.8 PH domain-assisted nucleotide exchange in LARG

Two contact regions between the LARG PH domain and RhoA could be responsible for PH-domain assisted nucleotide exchange in LARG. The first is a direct contact between the N-terminus of the αC helix of the LARG PH domain and the Glu^{97} side chain of RhoA. Specifically, Glu^{97} of RhoA forms hydrogen bonds with the side chain and backbone amide of Ser^{1118} of LARG. This interface buries about 200 $Å^2$ of surface area. The second region involves interactions between the DH domain and RhoA, which are supported by residues in the PH domain. The end of α6 of the DH domain (residues 975 to 986), which constitutes the hinge between the DH and PH domains that moves upon RhoA binding, has several interactions with RhoA-Arg^{68} and these interactions are supported by residues from the β1 strand of the PH domain. In particular, Glu^{982} and Asn^{983} of the α6 segment of LARG both coordinate the guanidino group of Arg^{68} of RhoA with a salt bridge and a hydrogen bond, respectively. The side chain of Arg^{68} is also sandwiched between Arg^{986} and Lys^{979}. All of these α6 residues are supported by direct and specific contacts with the PH domain. Asn^{983} is completely buried in the interface between the DH and PH domains and RhoA, and forms a hydrogen bond with the backbone carbonyl of Met^{1020} in β1 of the PH domain. Lys^{979} likewise forms a hydrogen bond with the backbone carbonyl of Ile^{1021}. Arg^{986} is supported by a salt bridge with Glu^{1023} in β1. Glu^{1023} also directly contacts Arg^{68} with a long-range salt bridge. Glu^{980} of α6, and Lys^{1019} of β1, form yet another specific contact that helps to cement the interface between DH and PH domains. These interdomain contacts are either

invariant or highly conserved among the Lbc family except for Glu^{1023}, which is substituted with serine in intersectin and cysteine in Net1 and Xpln.

The single direct contact between the PH domain and RhoA and the indirect contacts of the PH domain with RhoA at the DH-PH interface are reminiscent of what has been observed in the Dbs-Cdc42 [4] and Dbs-RhoA structures [3]. In both structures Dbs contacts the GTPase via a β3-β4 loop residue of the PH domain (Lys^{885}). However, mutation of contact residues on either the PH domain (Lys^{885}) or Cdc42 (His^{103}) did not influence exchange activity *in vitro*. On the other hand, a triad of residues from the PH domain (Tyr^{889}), the α6 segment of the DH domain (His^{814}) and Cdc42 (Asp^{65}) forms interactions, that seem to be important for nucleotide exchange, because a Y889F mutant has the most detrimental effect on nucleotide exchange. The equivalent triad is not observed in LARG. However, Lys^{979} of LARG is the equivalent residue to His^{814} in Dbs and when superimposed on RhoA, the Lys^{979}-Ile^{1021} contact in LARG is in an equivalent position to the His^{814}-Tyr^{889} contact in Dbs. Furthermore, Glu^{1023} of the PH domain hydrogen bonds to residues in the DH domain (Arg^{986}) as well as RhoA and might play an equally essential role to Tyr^{889} in Dbs.

3.3.9 A protein docking site on the LARG PH domain

In the DH/PH crystals, the PH domain forms a two-fold crystallographic dimer interface that buries 800 $Å^2$ of surface area (Fig. 6). The interface consists of a solvent-exposed hydrophobic patch on the β4-β7 sheet of the PH domain, which includes the side chains of Leu^{1086}, Ala^{1091}, Phe^{1098}, Ile^{1100}, Ala^{1107}, and Ile^{1109}. The hydrophobic patch of

all four of the PH domains in the asymmetric unit of the DH/PH-RhoA crystals form similar non-crystallographic dimer contacts. Furthermore, all four of the PH domains in the DH/PH-RhoA crystal asymmetric unit are similarly dimerized with translationally-related PH domains. Although the PH domains in other DH/PH structures are also involved in crystal contacts, an equivalent hydrophobic patch is not observed in the Dbs, Tiam or intersectin DH/PH domains. There is no evidence from size exclusion chromatography that the LARG DH/PH domains are dimeric in solution. However, the residues that compose this solvent-exposed patch are highly conserved among the Lbc family of RhoGEFs, suggesting a conserved functional role for these solvent-exposed residues. If the PH domain docks with the plasma membrane as it does in other well-characterized PH domains (*e.g.* PLCδ1 [109]), this patch would be able to interact laterally with another peripheral membrane protein or another domain. Based on these observations, the hydrophobic patch on the PH domain could represent a protein interaction site that could be exploited by either one of the other domains of LARG (*e.g.* PDZ or RH) or another protein (*e.g.* $G\alpha_{13}$).

3.3.10 Specificity of LARG for RhoA

RhoGEFs exhibit a wide range of substrate specificities, from promiscuous RhoGEFs like Vav, which catalyzes nucleotide exchange on all three Rho subfamily proteins to strictly specific RhoGEFs like LARG, which catalyzes nucleotide exchange on RhoA only [54]. The structure of LARG is the first of a RhoA-selective RhoGEF and in comparison with previously determined DH/PH domains-GTPase complexes of Rac1-

specific (Tiam), Cdc42-specific (intersectin) and Cdc42- and RhoA-specific (Dbs) GEFs, the structural basis of strict RhoA specificity can now be explained.

In previously studied RhoGEFs, the regions of the DH domain most important for determining substrate specificity were proposed to be the α4-α5 loop, which interacts with residues of the β1 strand of the GTPase, and the N-terminal segment of the α5 helix, which accommodates Trp^{58} of RhoA (Trp^{56} in Rac1, Phe^{56} in Cdc42). LARG indeed appears to use both of these regions to dictate specificity. The side chain of Trp^{58} of RhoA is completely buried in the LARG interface and forms a hydrogen bond with the carboxylate of LARG-Asp^{928} (Fig. 3.23) a residue conserved as aspartate or glutamate among Lbc subfamily RhoGEFs except intersectin, which has a serine at the equivalent position. Dbs, which is selective for either Cdc42 or RhoA, also has a serine (Ser^{763}) in this position, and accommodates either Trp^{58} of RhoA or the equivalent Phe^{56} of Cdc42 in part by altering the side chain conformation of Ser^{763} as well as of Leu^{766}. The equivalent residue to Dbs-Leu^{766} is LARG-Pro^{931}, whose side chain is relatively short and rigid. Therefore, the binding of Cdc42 to LARG is disfavored in part because its Phe^{56} residue would clash with the LARG-Asp^{928} side chain and leave a cavity in the interface next to LARG-Pro^{931}. However, LARG-Pro^{931} is not conserved throughout the RhoA-specific Lbc family RhoGEFs, but is a leucine in most members of this family. It can therefore not solely be responsible for RhoA specificity.

Although the region in the α5 helix, that accommodates Trp^{58}/Phe^{56} seems to play an essential role in the discrimination between RhoA and Cdc42, or Rac1 and Cdc42, it does not explain the discrimination between RhoA and Rac1, which both have a

tryptophan in this position. The second proposed specificity-determining contact region in the α4-α5 loop of the DH domain therefore has to be considered. The α4-α5 loop makes contacts to the β2-β3 sheet of the GTPase as well as to the N-terminal region of the GTPase. Arg^{923} in the α4-α5 loop of LARG forms salt-bridges with both Asp^{45} and Glu^{54} of the β2-β3 sheet of RhoA, which are substituted by shorter polar side chains in Rac1 and Cdc42 (Asn^{43}, Thr^{52}). Arg^{923} is invariant throughout the Lbc subfamily of RhoGEFs except intersectin, whose equivalent residue is glycine. Intersectin is the only Lbc subfamily RhoGEF that catalyzes nucleotide exchange on Cdc42 and not RhoA. The equivalent residue to LARG-Arg^{923} in Dbs is Lys^{758}, which has a hydrogen bond to RhoA-Glu^{54} in the Dbs-RhoA structure but no interactions with this region in the Dbs-Cdc42 structure. In Tiam, which is specific for Rac1, the equivalent residue is Gln^{1177}, which also does not interact with the β2-β3 sheet in Rac1. It therefore seems that contacts to the β2-β3 sheet are relevant for the specific recognition of RhoA, but do not play a role in the recognition of Cdc42 and Rac1. Contacts between the α4-α5 loop of the LARG DH domain and the N-terminal region of the GTPase also seem to be essential in the discrimination between RhoA and Rac1. Arg^{5} packs against the α4-α5 loop and forms a hydrogen bond with the Arg^{923} backbone carbonyl. Arg^{5} is substituted with an alanine in Rac1.

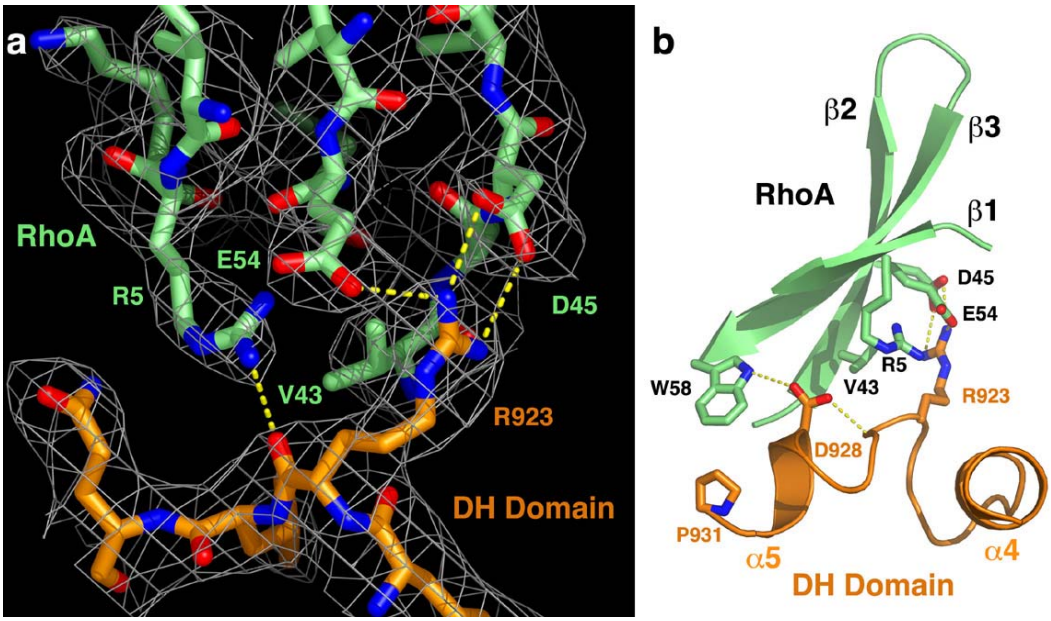

Figure 3.23: Contacts between the LARG DH/PH domains and RhoA expected to dictate substrate specificity. **(a)** As predicted in prior studies [3], Arg^{923} in the α4-α5 loop of LARG makes multiple salt bridge interactions with an acidic patch on the surface of RhoA that consists of Asp^{45} and Glu^{54} (substituted by asparagine and threonine, respectively, in both Rac and Cdc42). Arg^{5} of RhoA (substituted by alanine and threonine in Rac and Cdc42, respectively) packs against the α4-α5 loop of LARG and forms a hydrogen bond with the backbone carbonyl of Arg^{923}. Adjacent to Arg^{5}, a salt bridge is formed between LARG-Lys^{899} and RhoA-Asp^{76} (not shown). Asp^{76} is substituted by a glutamine in both Rac1 and Cdc42, and therefore this contact could also contribute to substrate specificity. The mesh cage represents a σ_A-weighted $2|F_o|-|F_c|$ Fourier map contoured at 1.0 σ, generated using the CCP4 program suite. **(b)** The same interface rotated 90˚ around a vertical axis, highlighting interactions with RhoA-Trp^{58}, which is completely buried at the interface. Asp^{928} of LARG, which is conserved as an acidic residue in all Lbc subfamily RhoGEFs except for intersectin, forms hydrogen bonds that bridge Trp^{58} of RhoA and a backbone nitrogen in the LARG α4-α5 loop. In Cdc42, Trp^{58} is substituted by phenylalanine, which would be incompatible with Asp^{928} of LARG and leave a cavity in the interface. Val^{43} of RhoA (substituted by serine and alanine in Rac and Cdc42, respectively) packs snuggly against the side chain of LARG-Arg^{923}. Conservative substitution of this valine residue with isoleucine, as in RhoC, could not be easily accommodated. Indeed, it was recently reported that Xpln, another Lbc subfamily RhoGEF selective for RhoA, can catalyze nucleotide exchange on RhoA and RhoB, but not RhoC.

4. Kinetic assessment of LARG fragments and mutants by fluorescence spectroscopy

4.1 Introduction

In order to confirm the functionality of the crystallized DH/PH domains, and to assess mutants, that were generated, based on the structural analysis, as well as various LARG fragments, a fluorescence assay was developed. Unlike Ras, Rho GTPases have a tryptophan in proximity to the nucleotide-binding site (Trp^{99} in RhoA). Therefore, various approaches can be taken to the kinetic assessment of RhoGEFs (Fig. 4.1). Previous fluorescence assays of DH/PH domains measured either fluorescence from the nucleotide-sensitive tryptophan (Fig. 4.1a) or from a fluorescent nucleotide (methylanthraniloyl (mant)-GTP or mant-GDP) (Figure 4.1b, c). For Ras and Ran it has been shown that mant-nucleotides bind with similar affinity compared to un-modified nucleotides and that the on and off rates are comparable [114, 116, 117].

Here, an assay was developed that records the decrease of fluorescence resonance energy transfer (FRET) from Trp^{99} in RhoA to mant-GDP upon its exchange for GTP (Figure 4.1d). This assay was developed with the intention of getting an improved signal to noise ratio. The assay was successfully used to assess the relative activities of various LARG fragments as well as of DH/PH domains mutants and RhoA mutants. Preliminary data for the determination of Michaelis-Menten parameters were also collected.

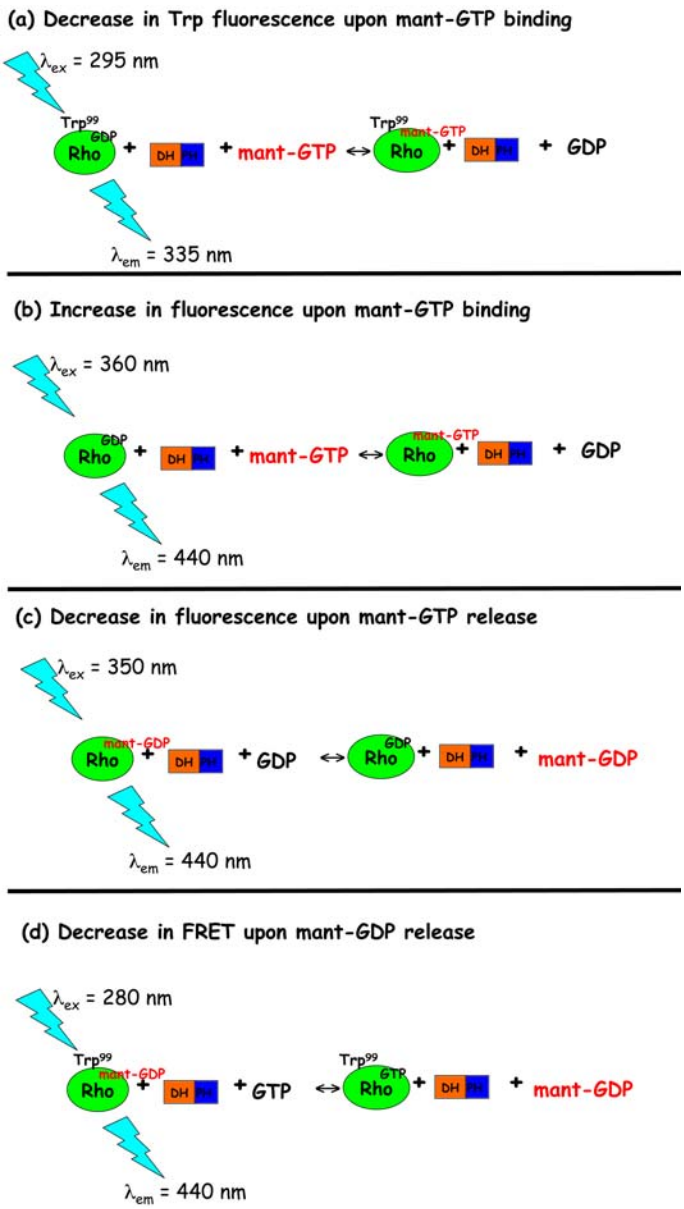

Figure 4.1: Various approaches for the assessment of the catalytic rates of RhoGEFs. (a) The decrease in tryptophan fluorescence upon exchange of GDP for a fluorescent nucleotide (methylanthraniloyl (mant) GTP) can be measured [4]. **(b)** The increase in fluorescence upon exchange of GDP for either mant-GTP or mant-GDP can be measured [3, 5]. **(c)** The decrease in fluorescence upon exchange of mant-GDP for GDP can be measured [7]. **(d)** The decrease in FRET upon exchange of GTP for mant-GDP has been measured in this study.

4.2 Methods

4.2.1 Cloning and Purification of proteins used in the fluorescence assays

A fragment of human LARG encoding the RH/DH/PH domains (residues 341 to 1138) was inserted into the pMALc2T-H_6 vector (see 3.2.1) using *BamH* I and *Sal* I restriction sites and purified as described for the DH/PH domains (see 3.2.1, Fig 4.2). Fragments encoding the LARG DH domain (residues 765 to 986) and RH/DH domains (residues 341 to 986) were cloned into the pMALc2H_{10}T vector (see 3.2.3) using *BamH* I and *Sal* I restriction sites and purified as described for RhoA (see 3.2.3, Figs 4.3, 4.4,), substituting buffers used for RhoA purification with buffers used for DH/PH domains purification. The yield for the RH/DH and RH/DH/PH fragments was 5 mgl/liter cell cultures. The yield for the DH domain was 15 mg/liter of cell culture.

Human Rac1 and Cdc42 were cloned from pCDNA3 vectors (a gift from Dr. S. Dharmawardhane) into the pMALc2H_{10}T plasmid and purified as described for RhoA (see 3.2.3).

DH/PH and RhoA mutants were generated by site directed mutagenesis using the Quickchange site directed mutagenesis kit (Stratagene) and purified as described for the wild type proteins (see 3.2.1, 3.2.3) [92].

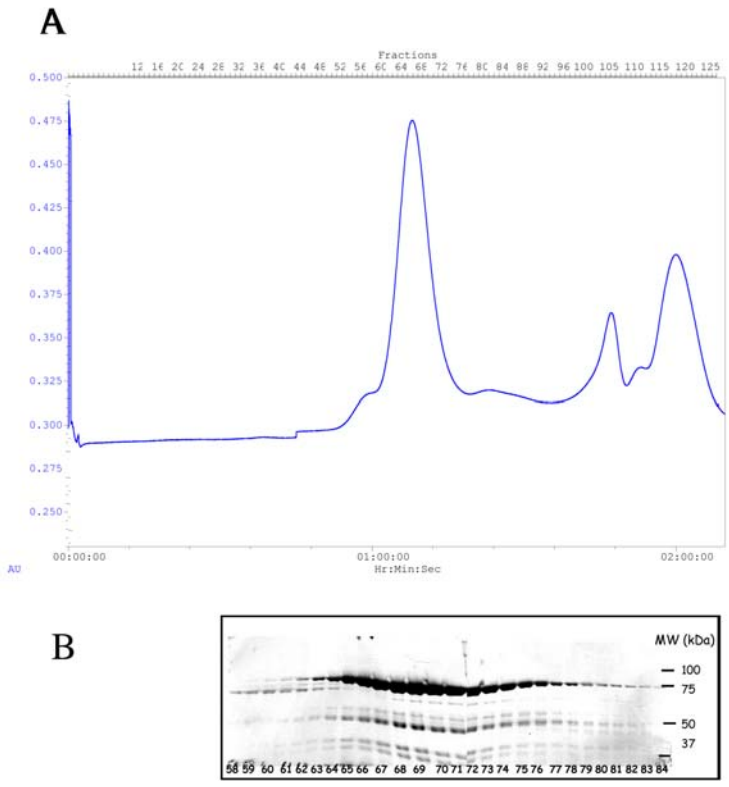

Figure 4.2: Purification of the LARG RH/DH domains. The LARG RH/DH domains were expressed from a modified pMAL vector as an N-terminal fusion protein with MBP and a C-terminal H_6 tag. The MBP-RH/DH-H_6 fusion protein was purified on a NiNTA column, digetsted with TEV protease, and the RH/DH-H_6 domains were separated from MBP by another NiNTA column. Subsequently the RH/DH-H_6 domains were purified on a gel filtration column (S200, 16/60, Amersham Pharmacia). The RH/DH-H_6 domains eluted at a flow rate of 1ml/min. Fractions 66-78 were pooled and concentrated to about 10 mg/ml.

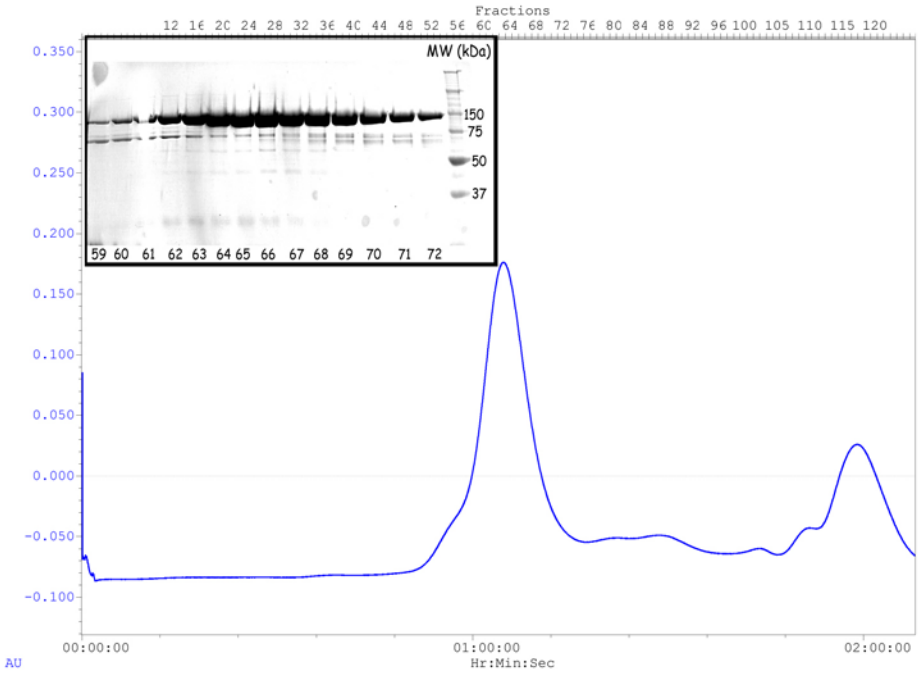

Figure 4.3: Purification of the LARG RH/DH/PH domains. The LARG RH/DH/PHdomains were expressed from a modified pMAL vector as an N-terminal fusion protein with MBP and a C-terminal H_6 tag. The MBP-RH/DH/PH-H_6 fusion protein was purified on a NiNTA column, digetsted with TEV protease, and the RH/DH/PH-H_6 domains were separated from MBP by another NiNTA column. Subsequently the RH/DH/PH-H_6 domains were purified on a gel filtration column (S200, 16/60, Amersham Pharmacia). The RH/DH/PH-H_6 domains eluted at a flow rate of 1ml/min. Fractions 62-72 were pooled and concentrated to about 10 mg/ml.

4.2.2 Fluorescence Assays

As shown in figure 4.1d GTPases were pre-loaded with mant-GDP and the exchange reaction was performed in the presence of the GEF and an excess of GTP. Nucleotide exchange was measured at an excitation wavelength of 280 nm and emission wavelength of 430 nm. At 280 nm a tryptophan (Trp^{99}) close to the nucleotide binding pocket is excited and loss of FRET from Trp^{99} to the mant group of mant-GDP is measured upon exchange of mant-GDP for GTP.

Loading of Rho GTPases with mant-GDP: GTPases were loaded with mant-GDP by incubating 100 to 200 µl of a 180 µM GTPase in loading buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 4 mM EDTA, 1 mM DTT) for 1.5 hrs on ice with a ten fold molar excess of mant-GDP (Jena Bioscience). Subsequently, mant-GDP loaded GTPases were stabilized by addition of $MgCl_2$ to a final concentration of 10 mM and incubation for an additional 30 min on ice. EDTA and free nucleotide were removed by exchanging mant-GDP loaded GTPases into reaction buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT) via gel filtration on a Sephadex G-10 (Sigma) column of approximately 3 ml bed volume. The gel filtration column was pre-equilibrated with reaction buffer. Mant-GDP loaded GTPase was eluted in approximately 100 µl fractions, and the protein content of the fractions was assessed by mixing 3 to 6 µl of the eluted fractions with 50 µl of Bradford solution. Mant-GDP loaded GTPases eluted between fractions 11 to 17. These fractions were pooled and the concentration of mant-GDP-RhoA was determined by absorption readings at 280 nm and at 360 nm, which is the

excitation maximum of mant-GDP. The extinction coefficient of mant-GDP is 5700 M^{-1} cm^{-1}. The ratio between concentrations determined at 360 nm and 280 nm gives an estimate of the ratio of mant-GDP loaded RhoA to total RhoA (loaded and unloaded RhoA). According to this estimation 50-90% loading was achieved using this method. Loading of Rac1 with mant-GDP was verified with the N-terminal DH/PH domains of Trio (a gift from Dr. Nicolas Nassar, Stony Brook University), that specifically exchanges nucleotides on Rac1 and RhoG, but not Cdc42 or RhoA [118, 119].

Nucleotide exchange assays: Exchange assays were performed on a FluoroMax-3 spectrofluorometer at 25° C, in a 200 μl cuvette (λ_{ex} = 280 nm, λ_{em} = 430 nm, slits = 2/2 nm). Ten μl of 2 mM GTP (final GTP concentration is 100 μM) were added to 180 μl of 1.1 μM mant-GDP loaded GTPases (final GTPase concentration is 1.0 μM) and the solution was subjected to two pulses of vortexing. The exchange reaction was started with the addition of 10 μl of a 2 μM GEF solution (final GEF concentration is 100 nM) and traces were recorded after two pulses of vortexing. Mant-GDP loaded GTPase was kept at 25° C, GTP and GEF solutions were kept on ice. Reactions were traced for 300 to 500 sec. The baseline of the uncatalyzed nucleotide exchange was traced for 1000 sec. Reactions were performed at least in triplicate.

Data assessment: k_{obs} were determined by fitting the traces to an equation describing a one-phase exponential decay using the program Graphpad Prism version 4.0. To obtain a better fit for the very flat traces of the uncatalyzed exchange the value for the total

fluorescence change (equivalent to the PLATEAU in the program Prism) was fixed to a value that corresonds to the total fluorescence change for the DH/PH domains catalyzed reaction.

4.3 Results and discussion

4.3.1 FRET *versus* mant-nucleotide and tryptophan fluorescence

Previous kinetic studies of RhoGEF DH/PH domains and their mutants were performed by either measuring changes in tryptophan fluorescence or mant-nucleotide fluorescence upon nucleotide exchange (Fig. 4.1a-c) [4, 5, 7]. Here we developed a FRET-based assay to increase the signal to noise ratio and to decrease the necessary amount of mant-nucleotide used (Fig. 4.1d). Although one of the previously used assays is based on the decrease of tryptophan fluorescence due to binding of and energy transfter to mant-nucleotide (Fig. 4.1a) [4], our assay would be expected to result in a higher signal to noise ratio because there is a larger difference between the excitation and emission wavelengths (Fig. 1.4). Indeed, only a two fold increase in signal to noise ratio was reported for Rap2A when the decrease of tryptophan fluorescence was measured [120]. In our assay, however, a 10 fold increase of the signal to noise ratio was observed, with FRET as opposed to mant-fluorescence, as determined by the maximal achievable fluorescence change of the GEF catalyzed reaction. Furthermore, a second tryptophan in RhoA (Trp^{58}) is intimately involved in the association with the DH domain and could contribute to the measured fluorescence changes of an exchange reaction. This contribution would shift the detected rates towards rates of binding rather than rates of nucleotide exchange. By measuring FRET these effects should be diminished because Trp^{58} is about 4 Å further away from the nucleotide than Trp^{99} and energy transfer decreases with the distance to the sixth power (r^6). Besides improving the signal to noise ratio, the assay developed here also reduces the amount of fluorescent nucleotide used in

the reaction mix, which is not only more economical but also decreases the possibility of internal fluorescence effects, due to the mant-nucleotide. At high concentrations of mant-nucleotide the fluorescence from the mant-group could be quenched by the nucleotide.

4.3.2 Fluorescence assays of DH/PH and RhoA mutants

Based on the structural analysis of the LARG DH/PH domains in complex with RhoA, mutants of the DH/PH domains and RhoA were designed to better understand the role of the unusual N-terminal subdomain of the LARG DH domain and the functional basis for PH domain-assisted nucleotide exchange in LARG. The following hypotheses were tested:

Hypothesis 1: The N-terminal extension is involved in nucleotide exchange. This might be due to interactions with switch I, which leave the nucleotide binding site more exposed and stabilize nucleotide-free RhoA. Deleting or disturbing the structural integrity of the N-terminal subdomain will result in decreased nucleotide exchange rates. Mutants that were generated to test this hypothesis are: ΔN-DHPH, DHPH-769A, DHPH-769D, DHPH-790G, ΔN-DHPH/790G, RhoA-33E, RhoA-27T.

Hypothesis 2: The N-terminal extension is a specificity-determining region for RhoA selective RhoGEFs. N-terminal deletion mutants will catalyze nucleotide exchange on Cdc42 and/or Rac1.

Hypothesis 3: PH domain-assisted nucleotide exchange in LARG is based on either the direct interaction of LARG-Ser^{1118} with RhoA-Glu^{97} and/or LARG-Glu^{1023} with RhoA-Arg^{68}. Allternatively, LARG-Glu^{1023} could indirectly act on RhoA by supporting

an interaction of the DH domain with RhoA (LARG-Arg^{986}-RhoA-Arg^{68}). Point mutations in this region on RhoA or the PH domain will therefore result in nucleotide exchange rates that are comparable to the DH domain alone. Mutants that were generated to test this hypothesis are: RhoA-97A, DHPH-1023A, DHPH-1023R.

Role of the N-terminal subdomain in nucleotide exchange – Hypothesis 1

Structural analysis of the DH/PH domains in complex with RhoA suggests a role for the N-terminal subdomain in RhoA binding and nucleotide exchange. The contribution of the N-terminal extension to nucleotide exchange activity was therefore investigated by assaying an N-terminal deletion mutant (ΔN-DHPH) as well as mutants that would be expected to exhibit problems with the structural integrity of the N-terminal subdomain (DHPH-769A, DHPH-769D, DHPH-790G). Furthermore, the residues in RhoA, that interact with the N-terminal subdomain have been mutated (RhoA 27T, RhoA 33E).

The Trp^{769} mutants reduced nucleotide exchange activity to about the same level as the N-terminal deletion mutant (13-20% of the wild type DH/PH domains, Fig 4.2). This suggests that Trp^{769} is important for the structural integrity of the N-terminal subdomain. The DHPH-790G mutant also reduced the catalytic activity (about 25% of wild type DH/PH, Fig 4.2), suggesting that this residue also contributes to the structural integrity of the subdomain. This result could also indicate that the substitution of Glu^{790} for glycine in intersectin, would not support an equivalent N-terminal subdomain.

Surprisingly, the RhoA 27T and RhoA 33E mutants were not compromised in their nucleotide exchange activities (Table 4.1). These results could be explained by the fact that the electron density for the RhoA-Lys^{27} side chain is weak, and this residue is perhaps unimportant for binding LARG. Val^{33} is partially solvent exposed, and substitution by a bulkier side chain could therefore be accommodated without compromising binding.

An interesting observation is that the Val^{33} mutant partially rescued the activity of the N-terminal deletion mutant (Fig 4.2). The reason for this observation is unclear. It might be possible that a residue in the α1 helix is re-oriented upon deletion of the N-terminus and favorably interacts with Glu^{33}. The only residue that could possibly be involved is Lys^{787}. However, it seems structurally very unlikely.

Given that the RhoA-33E mutation has no effect on nucleotide exchange, it is somewhat surprising that the DHPH-769A, DHPH-769D and ΔN-DHPH mutations of LARG had such a dramatic effect on activity. One explanation could be indirect effects on the conformation of Glu^{790} of LARG, whose side chain is constrained via three hydrogen bonds with the N-terminal subdomain, two to backbone amides and one to the side chain of Gln^{770}. Upon binding RhoA, the side chain of Glu^{790} is completely buried in the RhoA interface and packs against the side chain of Tyr^{34} in the RhoA effector loop. Therefore, destabilization of the N-terminal subdomain by point mutation of its core hydrophobic residue, or its complete removal, could lead to release of the Glu^{790} side chain, allowing it to assume a more extended conformation that interferes with RhoA binding. To test this hypothesis, we constructed a ΔN-DHPH/790G mutation of the

LARG DH/PH domain. However, this double mutant did not rescue the activity of the ΔN-DHPH mutant. The role of the N-terminal subdomain in nucleotide exchange on RhoA is therefore not dependent on its ability to sequester an extended conformation of Glu^{790}. The results from the Glu^{790} mutants show that this residue is indispensable for nucleotide exchange and suggest that its interface with Tyr^{34} of switch I is a critical interaction.

Although the reduced catalytic activity of the DH/PH mutants could also be due to a population of molecules that are in an unfolded or not properly folded state, they correspond to activities of mutants of the Dbs DH/PH domains, which were reduced to about 10-16% [4, 121]. For one of the Dbs mutants (Y889F) a crystal structure was determined, which showed that its overall fold was not disturbed [4].

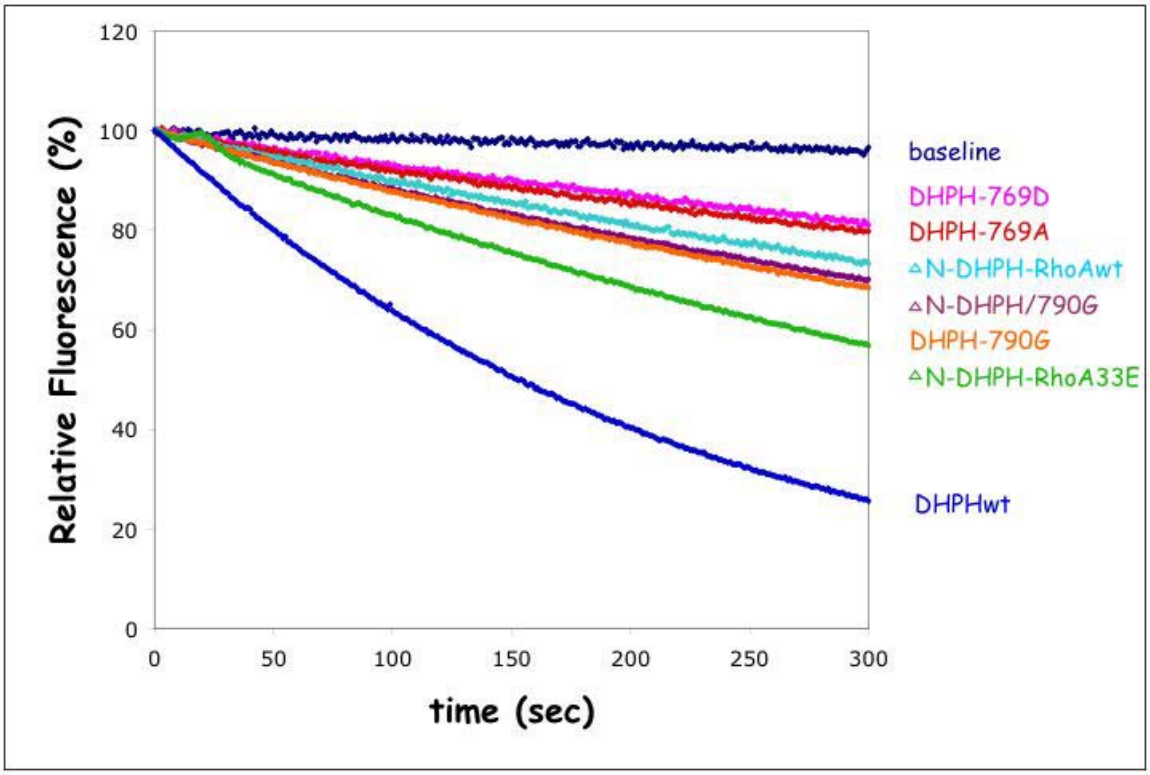

Figure 4.4: The N-terminal subdomain of the LARG DH/PH domains plays an essential role in nucleotide exchange. Nucleotide exchange on RhoA was measured using a FRET-based nucleotide exchange assay. For each time course, 1 μM of mant-GDP loaded RhoA was incubated with 100 μM GTP at 25 °C, and the exchange reaction was started by the addition of 100 nM LARG fragments or mutants, except for the baseline. The subsequent decrease in fluorescence (λ_{ex}=280 nm, λ_{em}=430 nm) was then measured for 300 sec. Each curve shown is the average of 2-3 measurements. Traces of baseline nucleotide exchange of the RhoA-33E mutant and traces of DH/PH domains catalyzed nucleotide on RhoA-33E are as for the RhoAwt and are not shown. The W769, E790, and the N-terminal deletion mutants are all predicted to interfere with the structural integrity of the N-terminal subdomain and all show similar levels of activity.

	Rate of Nucleotide Exchange ($s^{-1}x10^{-3}$) [‡]			
LARG Fragment	RhoA	RhoA-V33E	RhoA-K27T	RhoA-E97A
None	0.15 ± 0.007	0.13 ± 0.007	ND	ND
DH (765-986)	3.7 ± 0.25 (25)[†]	ND[‖]	ND	ND
DHPH (765-1138)	5.0 ± 0.25 (33)	5.5 ± 0.06 (42)	5.1 ± 0.84	5.6 ± 0.29
DHPH-ΔN	1.0 ± 0.07 (6.6)	1.9 ± 0.06 (15)	ND	ND
DHPH-W769A	0.75 ± 0.02 (5)	1.1 ± 0.07 (8.5)	ND	ND
DHPH-W769D	0.68 ± 0.02 (4.5)	1.0 ± 0.04 (7.7)	ND	ND
DHPH-E1023A	3.6 ± 0.1 (24)	ND	ND	ND
DHPH-E1023R	2.2 ± 0.04 (15)	ND	ND	ND
DHPH-E790G	1.3 ± 0.06 (8.6)	ND	ND	ND
DHPH-E790G/ΔN	1.2 ± 0.02 (8)	ND	ND	ND
RHDH (341-986)	1.49 ± 0.13 (12.1)	ND	ND	ND
RHDHPH (341-1138)	1.42 ± 0.11 (10.7)	ND	ND	ND

[‡]To calculate k_{obs}, each curve was fit to a single-order exponential decay with Prism v 4.0. Each data point is the average of at least three measurements.[†]
[\]Numbers in parentheses correspond to the fold rate enhancement over the basal rate in the absence of LARG, if measured.
[‖]ND: not determined.

Table 4.1: Summary of k_{obs} from wild-type and mutant LARG and RhoA proteins.

Role of the N-terminal subdomain in specific nucleotide exchange on RhoA – Hypothesis 2

Sequence comparison suggests that the N-terminal extension might play a role in specific recognition of RhoA, because amino acids that comprise the hydrophobic core of the N-terminal subdomain are only conserved throughout the Lbc family of RhoA specific Rho GEFs. The only exception is intersectin, which is the only family member that is specific for Cdc42. In support of this hypothesis, an interesting observation was made when an LARG DH/PH domain construct (residues 785-1140), that lacked the N-terminal extension, was studied [54]. This construct showed some nucleotide exchange on Cdc42. We therefore tested the ΔN-DHPH mutant (residues 782-1138) and the ΔN-DHPH/790G double mutant against Cdc42 and Rac1. The ΔN-DHPH/790G mutant was investigated to rule out the possiblility that a lack of activity it is not the result of interference of Glu^{790} with Cdc42 and/or Rac1 binding. However, both mutants did not show any exchange activity on Cdc42 and Rac1 (data not shown).

PH domain-assisted nucleotide exchange in LARG – Hypothesis 3

The invariant presence of a PH domain C-terminal to the DH domain in RhoGEFs suggests a direct role of the PH domain in catalysis. Indeed, increased *in vitro* exchange activities of DH/PH vs. DH domains have been reported for Dbs (26 fold) [4], Trio (2 fold) [122], and a fragment of the LARG DH/PH domains (residues 785-1140), that does not contain the N-terminal extension (values of the fold enhancement were not calculated for this protein) [54]. Consistent with these studies, the fragment of the LARG DH/PH

domains (residues 765-1138) that was used in crystallization showed about 2 fold higher activity than the DH domain. The fact that the Dbs and Trio DH domains are so much more compromised in their activities than the LARG DH domain compared to their respective DH/PH domains indicates that the PH domains of these proteins have stronger or more interactions with their GTPases. The β3-β4 loop of the Dbs PH domain indeed has extensive contacts with the switch II region of the GTPase and with the C-terminus of the α5 helix of the GTPase [4]. However, both interaction regions do not seem to be important for PH domain-assisted nucleotide exchange *in vitro*. An alternative explanation for the less compromised activity of the LARG DH domain compared to the DH/PH domains could be that the N-terminal subdomain of LARG contributes considerably to the binding of RhoA.

The only case where the basis of PH domain-assisted nucleotide exchange in a RhoGEF is well understood is in Dbs, where residues of the PH domain, that support interactions of the DH domain with the GTPase seem to be essential for catalysis (see 3.3.7) [4]. The LARG DH/PH-RhoA complex structure suggests a similar basis for PH domain-assisted nucleotide exchange. Therefore, the contribution of the contact of Glu^{1023} with RhoA (Arg^{68}) and the DH domain (Arg^{986}) was analyzed by site directed mutagenesis. Furthermore, the direct contact between Ser^{1118} and Glu^{97} was analyzed by investigating a RhoA-97A mutant. Nucleotide exchange of RhoA-97A was nearly identical to that of wild-type RhoA (data not shown), suggesting, that this small, unconserved yet direct interface between the PH domain of LARG and RhoA is not important for nucleotide exchange, at least *in vitro*. This notion is supported by the fact

that Ser^{1118} and adjacent residues are not conserved among either the Lbc subfamily or RH-RhoGEFs. A smilar, but even more surprising obervation was made in Dbs, where mutating both the Cdc42 (Arg^{66}) and the PH domain contact residue (Tyr^{889}), revealed that the Y889F mutant severely diminished nucleotide exchange activity even though the R66A mutant had the same activity as wild type RhoA [4]. Because Tyr^{889} supports contacts of the DH domain with Cdc42 (His^{814}-Cdc42-Arg^{66}), this result could suggest that, although the hinge region of the DH/PH domains can accommodate different side chains within the contact region on the GTPase, the full catalytic activity is dependent on the structural integrity of the DH/PH interface. However, the structure of the Dbs-Y889F mutant in complex with Cdc42 has been solved and the relative disposition of PH to DH domains is unchanged compared to wtDbs. Dbs catalyzed nucleotide exchanges in Dbs has been shown to be dependent on the interaction of Dbs-His^{814} with Cdc42-Asp^{65}, because mutations in either residue severly diminishes exchange activity. It was therefore suggested that Tyr^{899} stabilizes the electronic configuration of His^{814} to promote interaction with Asp^{65} in Cdc42.

The lack of an effect of the RhoA97A mutant in LARG catalysis *in vitro* does not rule out the possibility that this direct interaction between the LARG PH domain and RhoA is used for PH domain-assisted nucleotide exchange *in vivo*. For example, the mutation of a residue in the Dbs PH domain that directly contacts the GTPase (Lys^{885}), did not compromise activity *in vitro* [4]. However, in *in vivo* focus formation assays, this contact was also shown to be important [88]. These results suggest that *in vivo* the RhoGEF could be regulated through interactions of its PH domain with other

intramolecular domains (*e.g.* PDZ, RH domains) or other proteins (*e.g.* $G\alpha_{13}$, actin) or through interactions with membranes. *In vivo* PH domains might be positioned more optimally with respect to their adjacent DH domains and their respective GTPases, and disruption of this positioning *in vivo* might result in a more pronounced effect than *in vitro*. In *in vitro* assays however, the basal GEF activity might already be severly compromised compared to *in vivo* activities and mutations in the PH domain-RhoA interface will not be detected.

On the other hand, the E1023A mutant of LARG demonstrates diminished GEF activity towards RhoA to about the same extent as the DH domain alone (Fig 4.3). The E1023R mutation, which introduces electrostatic repulsion as well as potential steric clashes, is even less active. These results verify the importance of interactions between the DH and PH domains of LARG at the hinge region, and based on similarities with Dbs, suggest a general mechanism by which PH domains can assist nucleotide exchange in RhoGEFs.

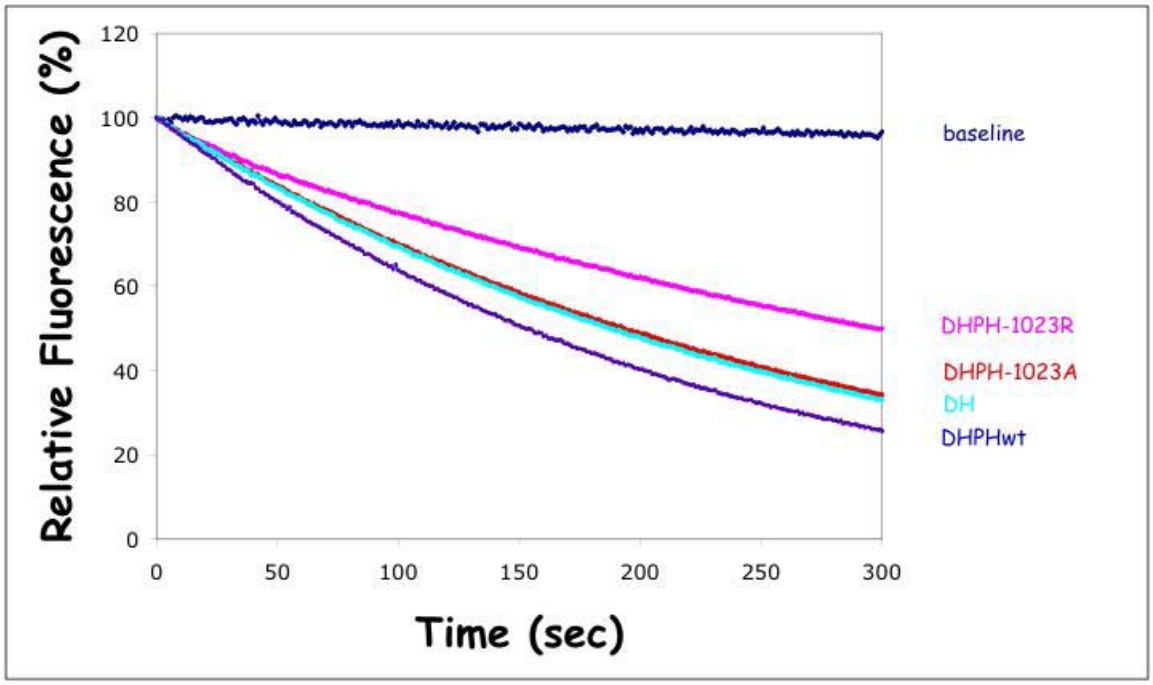

Figure 4.5: Mutation of Glu^{1023} eliminates PH domain-assisted nucleotide exchange in LARG. Nucleotide exchange on RhoA was measured using a FRET-based nucleotide exchange assay. For each time course, 1 μM of mant-GDP loaded RhoA was incubated with 100 μM GTP at 25 °C, and the exchange reaction was started by the addition of 100 nM LARG fragments or mutants, except for the baseline. The subsequent decrease in fluorescence (λ_{ex}=280 nm, λ_{em}=430 nm) was then measured for 300 sec. Each curve shown is the average of 2-3 measurements. Glu^{1023} has direct contacts to RhoA via a long-range salt bridge to Arg^{68} and indirect contacts via Arg^{986} in the DH domain. Mutations of Glu^{1023} reduce the activity of the DHPH domains to a level that is similar to the DH domain activity alone.

4.3.3 Fluorescence assays of LARG fragments containing the RH domain

A comparison of LARG constructs containing the RH domain with constructs without the RH domain shows an inhibitory role of the RH domain *in vitro*. This is consistent with initial data reported on p115RhoGEF catalyzed exchange on un-prenylated RhoA [56], that showed increased activity when the RH domain was truncated and would support an auto-inhibitory model for the effect of the LARG RH domain. However, when prenylated RhoA was used in the p115RhoGEF assays, a construct that lacked the RH domain showed decreased activity [57, 58]. Because data from assays with prenylated RhoA are not available for LARG, it can only be assumed from the close sequence homology between p115RhoGEF and LARG that the LARG RH domain also contributes to nucleotide exchange *in vivo*. However, there are differences in the p115RhoGEF constructs compared to the LARG constructs, which make it hard to apply conclusions from the p115RhoGEF assays to LARG. The p115RhoGEF DHPH construct contained approximately 100 amino acids N-terminal to the DH domain, which could affect the exchange rates. Furthermore, the results from our flourescence assays could, indeed represent the *in vivo* activity of the RH domain containing LARG fragments, if the presence of the PDZ domain in LARG results in a different domain architecture compared to p115RhoGEF, which does not have a PDZ domain.

It is curious that the LARG PH domain does not contribute to the exchange activity when the RH domain is present, instead the RHDHPH and RHDH constructs have comparable activities. This could mean that the RH domain associates with the DH/PH domains in a way that prevents the PH domain from assuming a favorable

relative position to the DH domain. These results are in contrast to the p115RhoGEF constructs where the PH domain also enhances nucleotide exchange in a construct that contains the RH domain (5 fold higher activity of RHDHPH compared to RHDH). However, in p115RhoGEF the DHPH construct has an about 14 fold higher activity compared to the DH construct. The RH domain of p115RhoGEF therefore also seems to diminish the effect of the PH domain on nucleotide exchange.

The fact that the N-terminal LARG mutants show approximately the same activity as the RH domain containing constructs might be an indication that the RH domain interacts with the N-terminal subdomain and effectively keeps it from eliciting its supporting role and therefore keeps DH/PH domains in a less active state. LARG activation by $G\alpha_{12/13}$ binding to the RH domain could change the conformation and release the N-terminus into a conformation that leads to full activity.

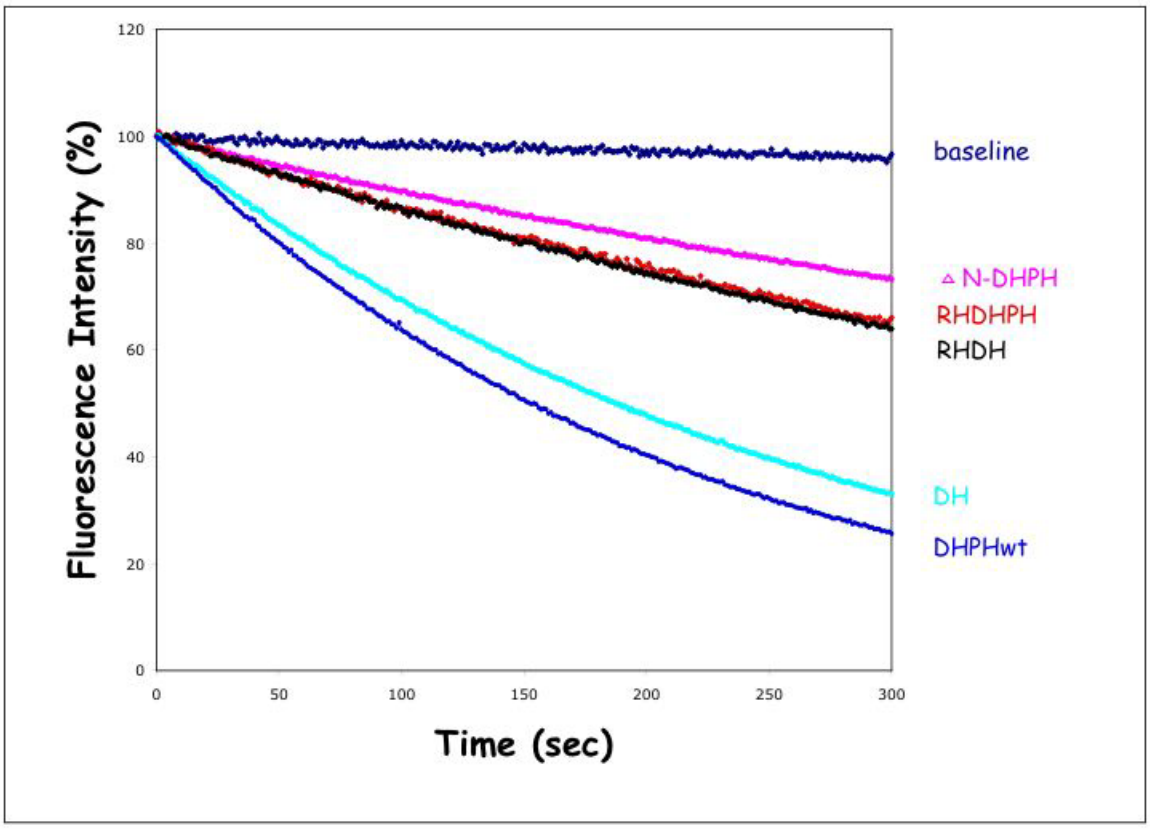

Figure 4.6: LARG constructs including the RH domain have diminished nucleotide exchange activity *in vitro*. Nucleotide exchange on RhoA was measured using a FRET-based nucleotide exchange assay. For each time course, 1 μM of mant-GDP loaded RhoA was incubated with 100 μM GTP at 25 °C, and the exchange reaction was started by the addition of 100 nM LARG fragments or mutants, except for the baseline. The subsequent decrease in fluorescence (λ_{ex}=280 nm, λ_{em}=430 nm) was then measured for 300 sec. Each curve shown is the average of 2-3 measurements. The fact that the N-terminal LARG mutants show approximately the same activity as the RH domain containing constructs might be an indication that the RH domain interacts with the N-terminal subdomain and effectively keeps it from eliciting its supporting role and therefore keeps DH/PH domains in a less active state. LARG activation by $G\alpha_{12/13}$ binding to the RH domain could change the conformation and release the N-terminus into a conformation that leads to full activity.

4.4 Preliminary determination of Michaelis-Menten parameters

Attempts were made to determine Michaelis-Menten parameters. Michaelis-Menten kinetics are considered steady state kinetics, following a scheme shown in Figure 4. The 'steady state' in Michaelis-Menten measurements refers to a condition where the concentration of the enzyme-substrate complex ([ES]) does not significantly change. Michaelis-Menten kinetics work under several assumptions. (1) The enzyme-substrate complex is in equilibrium with free enzyme and free substrate and therefore k_2 is much smaller than k_{-1}. (2) The enzyme is present in limiting concentrations relative to the substrate and all enzyme is present in a complex with the substrate. (3) The substrate concentration is much higher than the enzyme concentration and therefore does not significantly change during the enzyme-catalyzed reaction. Assuming Michaelis-Menten conditions K_M can be equated with K_D and an increase in K_M could therefore be an indication for impeded substrate binding. Changes in V_{max} on the other hand suggest changes at the catalytic site.

Michaelis-Menten scheme:

$$E + S \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} ES \overset{k_2}{\Rightarrow} E + P$$

Michaelis-Menten parameters:

$$v_0 = \frac{V_{max} \cdot S}{K_M + S} = \frac{E_{tot} \cdot k_{cat} \cdot S}{K_M + S}$$

$$K_M = \frac{k_{-1} + k_2}{k_1}$$

Assumptions of Michaelis-Menten kinetics:

(1) Equilibrium between E + S <--> ES, $k_{-1} >>> k_2$

(2) [E] is limiting --> $[ES] = [E]_{tot}$

(3) [S] >>> [E] --> $[S]_{free} = [S]_{tot}$

Figure 4.7: Summary of the assumptions in Michalis-Menten-type kinetics.

4.4.1 Methods

30 nM DH/PH domains were incubated with 0.5-4 μM mant-GDP loaded RhoA and the initial velocity was determined by fitting a line to the first 100 sec. An average of at least two measurements is reported. The initial velocity was plotted against the RhoA concentration and a hyperbola was fit to the curve using the program Graphpad Prism version 4.0.

4.4.2 Results and Discussion

The k_{cat} for the LARG DH/PH domains is comparable to k_{cat} values reported for the catalytic domain of SopE, a RhoGEF from *Salmonella enterica*, as well as the Dbl family RhoGEFs Dbl, and Lfc (Table 4.2). The increased K_M of the LARG DH/PH domains may be a result of the fact that K_M for Dbl and Lbc were determined with full length proteins and additional domains in these RhoGEFs might contribute to binding of the GTPase. The preliminary data for LARG are indicating that the developed FRET assay can be used for the determination of Michaelis-Menten parameters. However, the results have to be interpreted carefully because initial velocities were only determined over a small substrate concentration range. Furthermore, the RHDH and RHDHPH constructs were partially degraded during purification (Fig. 4.2 and 4.3) and the diminished GEF activities of these constructs could therefore be a result of a smaller fraction of active protein. Although activities of various RhoGEF DH/PH domains have been assessed *in vitro* using fluorescence and filter binding assays, Michaelis-Menten parameters have not been determined for most RhoGEFs. It is therefore impossible to compare the activity of the LARG DH/PH domains to activities that have been reported for other DH/PH domains of known structure. This would be interesting, because differences in k_{cat} and K_M could indicate if the N-terminal subdomain of the LARG DH/PDH domains contributes directly to the catalysis or to the binding of RhoA.

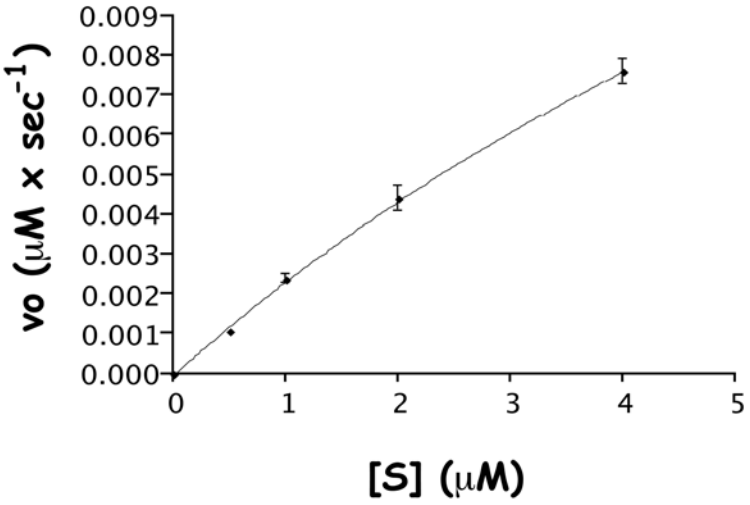

Figure 4.8: Preliminary Michalis-Menten plot for the LARG DH/PH domains. 30 nM DH/PH domains were incubated with 0.5-4 μM mant-GDP loaded RhoA and the initial velocity was determined by fitting a line to the first 100 sec. An average of at least two measurements is reported. The initial velocity was plotted against the RhoA concentration and a hyperbola was fit to the curve using the program Graphpad Prism version 4.0. The Michaelis-Menten parameters are comparable to the ones reported for other RhoGEFs (see Table 4.2).

	K_M (μM)	kcat (sec^{-1})
LARG (DH/PH)	**11.92 ± 1.8**	**1.01 ± 0.004**
Dbl (full length) [123]	**0.151 ± 0.003**	**1.98 ± 0.10**
Lbc (full length) [123]	**0.177 ± 0.03**	**1.16 ± 0.3**
SopE (catalytic domain) [124]	**1.7 ± 0.8**	**1.1 ± 0.1**

Table 4.2: Michaelis-Menten paramters of the LARG DH/PH domains in comparison to other RhoGEFs.

4.5 Technical problems with the fluorescence assays

Inconsistency of A_{280} readings compared to Bradford and BCA

Absorption readings at 280 nm consistently yielded lower concentrations for the DH domain and higher concentrations for the RHDHPH domains compared to the Bradford and bicinchoninic acid (BCA) methods. Concentrations were the same for the DHPH and RHDH domains. Protein concentrations for the determination of exchange rates were determined by the Bradford method. Protein concentrations for the assessment of Michaelis-Menten parameters were determined by absorption readings at 280 nm. The concentration of the DH domain, determined by absorption was consistently about half as much as when determined by the Bradford method. Although the DH domain concentration in these assays might therefore have been about twice as high, *i.e.* 10 nM instead of 5 nM, concentrations of the DHPH domains measured by absorption and Bradford always corresponded. Absorption readings at 280 nm detect tryptophan, tyrosine and cysteine residues. The Bradford reagent is thought to detect basic and aromatic residues, especially arginine [125], and the BCA reagent detects cysteine, tryptophan and tyrosine residues [126]. The DH domain contains 3 tryptophan, 5 tyrosine, and 4 cysteine residues. The DHPH domains contain 5 tryptophan, 9 tyrosine, and 6 cysteine residues. There is no consensus on what methods gives the most accurate data for protein concentration determination[75]. However, if the concentrations determined by 280 nm readings would be correct, our fluorescence assay would show that the DH domain has more activity than the DHPH domains and therefore that the PH domain has some inhibitory activity. This is in complete contradiction to what has been

found for other RhoGEF PH domains so far, with the exception of the SOS DH/PH domains, which do not show any activity *in vitro* [86]. It would also not be explained by structural analysis of the DHPH domains, where the PH domain does not interfere with RhoA binding.

Mant-GDP loaded GTPases get lost on re-used Sephadex G-10 gel filtration columns

Approximately 200 μM of RhoA were loaded with mant-GDP and the concentration after gel filtration was 60 to 100 μM for the first two loadings but only 25 and 10 μM for the subsequent loadings. These concentrations were determined by absorption readings at 280 nm. It might be possible that during the loading process a nucleotide free form of RhoA precipitates, binds to the gel filtration column and impedes subsequent gel filtration runs.

5. Conclusions and future studies

The structures of the LARG DH/PH domains alone and in complex with their substrate GTPase RhoA were determined. These structures show the same DH/PH domains in its GTPase-bound and -unbound form, and therefore, provide a dynamic view of the activation of a GTPase by a RhoGEF. The structure of the DH/PH domains of LARG is the first structure of RhoA selective DH/PH domains and the first of an RH-RhoGEF subfamily protein. The structure exhibits novel features (*i.e.* an αN1/αN2 subdomain, a hydrophobic patch on the PH domain, different conformation of the nucleotide binding site) that might be conserved in the two other RH-RhoGEF family members PDZ-RhoGEF and p115RhoGEF, which are all regulated by $G\alpha_{12/13}$ proteins. The functional relevance of the αN1/αN2 subdomain for nucleotide exchange activity has been demonstrated by fluorescence assays using N-terminal mutants of the DH/PH domains.

The role of the RH and PH domains has been investigated with the fluorescence assay, using various LARG fragments and mutants. The functional assays demonstrate PH domain-assisted nucleotide exchange in LARG and an inhibitory role of the RH domain. Results from structural and functional analyses taken together allow us to generate a model of LARG activation by $G\alpha_{12/13}$ at the cell membrane (Figure 5.1) or as part of cytoskeleton-attached signaling complexes, that could be applicable to the other RH-RhoGEF family members or the larger Lbc subfamily of RhoGEFs.

However, further structural and functional analyses will need to be performed to gain a complete understanding of LARG activity and regulation. Furthermore, the

development of *in vivo* assays or *in vitro* assays that better mimic the cellular environment in which LARG signaling takes place is inevitable for the understanding of LARG activity and regulation in the cell.

The LARG αN1/αN2 subdomain

The most distinct structural feature of the LARG DH/PH domains compared to other DH/PH domains is its N-terminal αN1/αN2 subdomain, which is seems to be important for nucleotide exchange activity, as determined by an *in vitro* fluorescence assay. An N-terminal deletion mutant decreased the nucleotide exchange activity to only 20 % of the wild type DH/PH domains. The αN1/αN2 subdomain of LARG directly contacts the RhoA switch I region, which contributes residues to the nucleotide binding pocket. However, it seems that this direct interaction is not critical for nucleotide exchange, but rather indirect interactions via the N-terminus of the α1 helix of the DH domain. Glu^{790} of the α1 helix is fixed in its conformation by a hydrogen bonding network with the N-terminal subdomain and packs against Tyr^{34} in the RhoA switch I region. E790G mutants retain only about 25% of the exchange activity compared to the wild type DH/PH domains. It might be possible that the N-terminal subdomain through this indirect interaction with switch I stabilizes the nucleotide-free form of RhoA. This could be either by supporting expulsion of nucleotide by opening the binding pocket and/or by relaying changes in the switch I region to other nucleotide binding regions. Although such a relay mechanism is not obvious from the structure, the LARG structure is the first of a DH/PH-GTPase complex, in which switch I resumes a more open

conformation, and where there is interference of the P-loop with the β-phosphate binding site and interference of GXKXDL and SAK motif residues with the purine binding site.

LARG activity could therefore be regulated via the N-terminal subdomain of its DH domain, which could be used as a switch mechanism. The N-terminal subdomain might be a target site for regulation by other proteins (e.g. $G\alpha_{13}$) or other domains within LARG (e.g. RH, PDZ domains). For instance the RH domain could sequester the N-terminal subdomain and therefore leave the DH/PH domains in a less active state. Interestingly, LARG constructs that also contain the RH domain (RHDH, RHDHPH) show levels of activity that are similar to an N-terminal truncated construct of the DH domain. It can therefore be envisioned that in the resting state of the cell, the RH domain binds to the αN1/αN2 subdomain of the DH/PH domains, sequestering it in its less active state, where Glu^{790} is positioned in a less favorable conformation for RhoA interaction. Upon activation of $G\alpha_{12/13}$ coupled receptors LARG might be recruited to the cell membrane by binding of the RH domain to $G\alpha_{12/13}$. This could in turn relieve the interaction of the RH domain with the αN1/αN2 subdomain of the DH/PH domains and lead to its activation.

Further structural and functional analyses will be necessary to elucidate the function of the αN1/αN2 subdomain and a possible role of the RH domain in regulating this subdomain. Mutants of RhoA-Tyr^{34} will have to be generated and assessed in the fluorescence assay to confirm the possible importance of the Glu^{790}-Tyr^{34} interface. A structure of the N-terminally truncated DH/PH domains in complex with RhoA might trap the DH/PH domains in its less active conformation and the purine and β-phophate

binding sites might be un-occupied like in other DH/PH-GTPase complexes, where an N-terminal subdomain is not present. A similar result might be seen in a structure of the RHDHPH construct in complex with RhoA. The comparison of K_M and k_{cat} values of the ΔN-DHPH and the wild type DH/PH constructs might be valuable in assessing if the subdomain only contributes to the binding of RhoA or also to the catalysis. Functional assays with prenylated RhoA and detergent micelles will be necessary to rule out similar contradictory results of RH domain containing constructs that were seen in p115RhoGEF when soluble RhoA was used [27].

PH domain assisted nucleotide exchange in LARG

The almost invariant position of a PH domain N-terminal to the DH domain in RhoGEFs suggests a role for the PH domain in nucleotide exchange. Although PH domain assisted nucleotide exchange *in vitro* has only been demonstrated for LARG, Dbs [4], and Trio [9], it is likely that most RhoGEFs show PH domain assisted nucleotide exchange under *in vivo* conditions. For example, the intersectin PH domain does not contribute to nucleotide exchange *in vitro*, but is necessary for full exchange *in vivo* [113]. This suggests that the PH domain could be a docking site for either regulatory proteins or for regulation by membranes or phospholipids. Indeed, the LARG structures reveal a hydrophobic patch that could be a potential protein docking site. Furthermore, the LARG PH domain has an electrostatic surface potential that is consistent with the surface potential of phospholipid binding PH domains, and a possible binding pocket for inositol phoshphates between the β1-β2 and β3-β4 loops, comprised of conserved

positively charged residues, is present. The orientation of the phospholipid binding site relative to the hydrophobic patch, would allow both, membrane association and interaction with a regulatory protein via interaction with the hydrophibic patch at the same time. It can therefore be envisioned that once LARG is recruited to the membrane via interaction of its PH domain with membrane located $G\alpha_{13}$ and/or phospholipid head groups the PH domain is positioned in a more optimal orientation relative to the DH domain and RhoA. In addition to the possible activation via release of the N-terminal subdomain (see above) this would provide a mechanism to even further enhance the activity of LARG (Figure 5.1). Another possibility is that the LARG PH domain binds to filamentous actin (F-actin). For instance the PH domain of Bruton's tyrosine kinase (Btk) binds to F-actin via a region that is overlapping with its inositol binding site [80]. The hydrophobic patch of LARG could at the same time bind to another actin binding protein. The possibility that the LARG PH domain is regulated by $G\alpha_{13}$ and/or phospholipid binding could be investigated by employing an adapted fluorescence assay with prenylated RhoA and lipid vesicles. Furthermore, binding studies of various LARG fragments with $G\alpha_{13}$ could confirm the proposed model and in the long term allow the crystallization of an LARG fragment in complex with $G\alpha_{13}$. These experiments require a source for the large scale production of $G\alpha_{13}$, which is currently not available (see Chapter 2).

A high degree of flexibility of the DH-PH junction has been demonstrated by our structures, where, upon complex formation, the PH domain rotates by about 30° relative to the DH domain to engage in RhoA binding. Two contact regions between the PH

domain and RhoA were identified in the complex structure. The first one is a direct contact between Ser^{1118} in the αCT helix of the PH domain and RhoA-Glu^{97}. The second contact region is an indirect contact in the hinge region between the DH and PH domains, where PH domain residues support contacts of the DH domain with RhoA-Arg^{68}. The importance of residues in the hinge region between the DH and PH domains has been demonstrated by the loss of PH domain-assisted nucleotide exchang in the Glu^{1023} mutants. Glu^{1023} is the central residue in a hyrdogen bonding network between DH domain residues (Phe^{990}, Arg^{986}), PH domain residues (Phe^{1039}) and RhoA residues (Arg^{68}). The direct contact between the PH domain and RhoA (LARG Ser^{1118}-RhoA Glu^{97}) does not seem to be important for nucleotide exchange *in vitro*, since a RhoA E97A mutant is an equally good substrate of LARG as wild type RhoA. However, this does not exclude the possibility that this direct contact contributes to PH domain-assisted nucleotide exchange *in vivo*, similarily to the direct PH domain-Cdc42 contact in the Dbs structure (Dbs L^{885}-Cdc42 His^{103}). A Dbs K885A mutant had unchanged exchange activity *in vitro* [4] but lost about 30% of its transforming ability [88]. To investigate the functional relevance of PH domain assisted nucleotide exchange in LARG it is therefore necessary to test the PH domain mutants in either *in vivo* assays or *in vitro* assays that better mimic an *in vivo* environment. The fluorescence assay described in Chapter 4 could be adapted to conditions that better mimic an *in vivo* environment by using prenylated RhoA and lipid vesicles. For p115RhoGEF it has been shown that prenylated RhoA is a much more potent substrate [27]. This might be due to either contribution of the prenyl group to protein-protein interactions, or the RhoGEF might be positioned on

the surface of the lipid vesicles in a manner that is more optimal relative to the orientation of prenylated RhoA on these vesicles. Using this adapted assay might allow the detection of the importance of the Ser^{1118}-Glu^{97} interaction. It would also be interesting to investigate a possible activation of the LARG DH/PH domains by phospholipid head groups, which could be used in the lipid vesicles. However, activation of nucleotide exchange on GTPases by phospholipids has been observed [127] and appropriate controls therefore have to be taken.

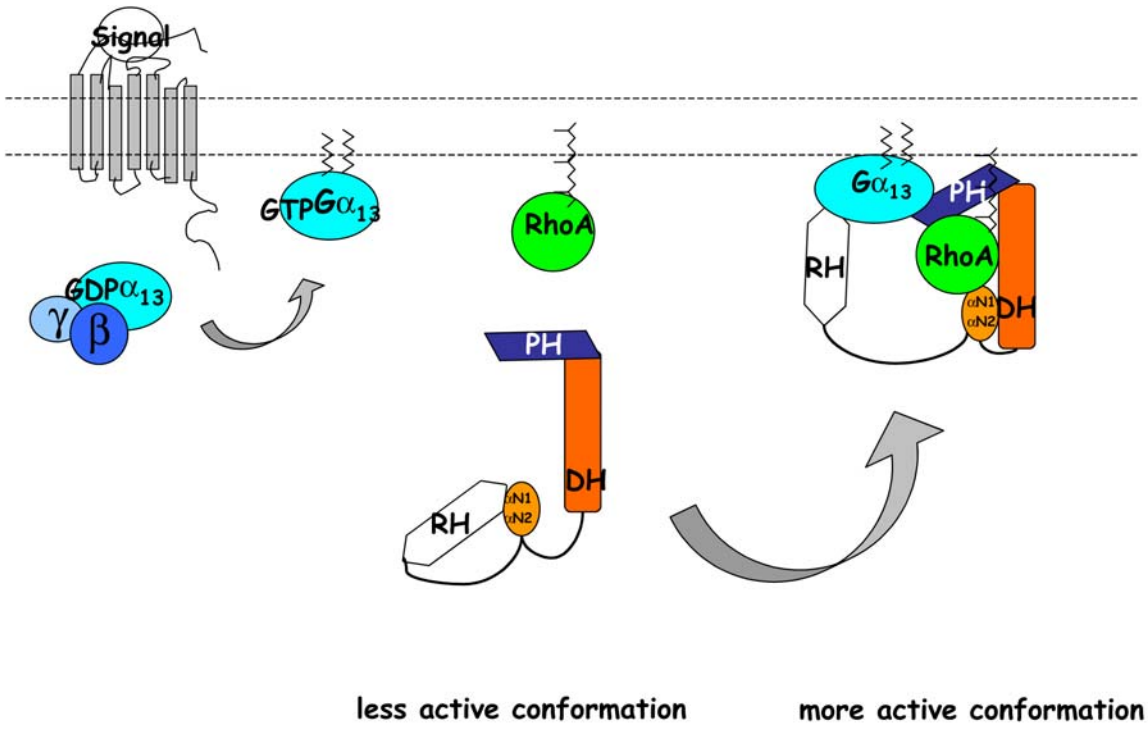

Figure 5.1: A model of signaling from LARG to RhoA, based on the structures of the LARG DH/PH domains and their complex with RhoA (see Chapter 3), as well as on functional assays of RhoA and LARG mutants (see Chapter 4). In the resting state of the cell the LARG DH/PH domains are in a less active conformation, due to sequestering of the αN1/αN2 subdomain by the RH domain. Activated $G\alpha_{13}$ could release this interaction by binding to the RH domain. This could result in the transformation to the more active conformation, where a more compatible surface for the binding of RhoA is available. At the same time the PH domain could assume a more compatible position for the binding of RhoA, due to its interaction with either $G\alpha_{13}$ or membrane phospholipids or both.

BIBLIOGRAPHY

1. Shimizu, T., K. Ihara, R. Maesaki, S. Kuroda, K. Kaibuchi, and T. Hakoshima, *An open conformation of switch I revealed by the crystal structure of a Mg2+-free form of RHOA complexed with GDP. Implications for the GDP/GTP exchange mechanism.* J Biol Chem, 2000. 275(24): p. 18311-7.
2. Boriack-Sjodin, P.A., S.M. Margarit, D. Bar-Sagi, and J. Kuriyan, *The structural basis of the activation of Ras by Sos.* Nature, 1998. 394(6691): p. 337-43.
3. Snyder, J.T., D.K. Worthylake, K.L. Rossman, L. Betts, W.M. Pruitt, D.P. Siderovski, C.J. Der, and J. Sondek, *Structural basis for the selective activation of Rho GTPases by Dbl exchange factors.* Nat Struct Biol, 2002. 9(6): p. 468-75.
4. Rossman, K.L., D.K. Worthylake, J.T. Snyder, D.P. Siderovski, S.L. Campbell, and J. Sondek, *A crystallographic view of interactions between Dbs and Cdc42: PH domain-assisted guanine nucleotide exchange.* Embo J, 2002. 21(6): p. 1315-26.
5. Worthylake, D.K., K.L. Rossman, and J. Sondek, *Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1.* Nature, 2000. 408(6813): p. 682-8.
6. Soisson, S.M., A.S. Nimnual, M. Uy, D. Bar-Sagi, and J. Kuriyan, *Crystal structure of the Dbl and pleckstrin homology domains from the human Son of sevenless protein.* Cell, 1998. 95(2): p. 259-68.
7. Aghazadeh, B., W.E. Lowry, X.Y. Huang, and M.K. Rosen, *Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation.* Cell, 2000. 102(5): p. 625-33.
8. Aghazadeh, B., K. Zhu, T.J. Kubiseski, G.A. Liu, T. Pawson, Y. Zheng, and M.K. Rosen, *Structure and mutagenesis of the Dbl homology domain.* Nat Struct Biol, 1998. 5(12): p. 1098-107.

9. Liu, X., H. Wang, M. Eberstadt, A. Schnuchel, E.T. Olejniczak, R.P. Meadows, J.M. Schkeryantz, D.A. Janowick, J.E. Harlan, E.A. Harris, D.E. Staunton, and S.W. Fesik, *NMR structure and mutagenesis of the N-terminal Dbl homology domain of the nucleotide exchange factor Trio.* Cell, 1998. 95(2): p. 269-77.
10. Buchwald, G., A. Friebel, J.E. Galan, W.D. Hardt, A. Wittinghofer, and K. Scheffzek, *Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE.* Embo J, 2002. 21(13): p. 3286-95.
11. Renault, L., J. Kuhlmann, A. Henkel, and A. Wittinghofer, *Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1).* Cell, 2001. 105(2): p. 245-55.
12. Goldberg, J., *Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching.* Cell, 1998. 95(2): p. 237-48.
13. Yu, H. and S.L. Schreiber, *Structure of guanine-nucleotide-exchange factor human Mss4 and identification of its Rab-interacting surface.* Nature, 1995. 376(6543): p. 788-91.
14. Kawashima, T., C. Berthet-Colominas, M. Wulff, S. Cusack, and R. Leberman, *The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution.* Nature, 1996. 379(6565): p. 511-8.
15. Palczewski, K., T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, and M. Miyano, *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. 289(5480): p. 739-45.
16. Lenzen, C., R.H. Cool, H. Prinz, J. Kuhlmann, and A. Wittinghofer, *Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25Mm.* Biochemistry, 1998. 37(20): p. 7420-30.

17. Wieden, H.J., K. Gromadski, D. Rodnin, and M.V. Rodnina, *Mechanism of elongation factor (EF)-Ts-catalyzed nucleotide exchange in EF-Tu. Contribution of contacts at the guanine base.* J Biol Chem, 2002. 277(8): p. 6032-6.
18. Gromadski, K.B., H.J. Wieden, and M.V. Rodnina, *Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu.* Biochemistry, 2002. 41(1): p. 162-9.
19. Romero, G., V. Chau, and R.L. Biltonen, *Kinetics and thermodynamics of the interaction of elongation factor Tu with elongation factor Ts, guanine nucleotides, and aminoacyl-tRNA.* J Biol Chem, 1985. 260(10): p. 6167-74.
20. Hart, M.J., A. Eva, D. Zangrilli, S.A. Aaronson, T. Evans, R.A. Cerione, and Y. Zheng, *Cellular transformation and guanine nucleotide exchange activity are catalyzed by a common domain on the dbl oncogene product.* J Biol Chem, 1994. 269(1): p. 62-5.
21. Klebe, C., H. Prinz, A. Wittinghofer, and R.S. Goody, *The kinetic mechanism of Ran--nucleotide exchange catalyzed by RCC1.* Biochemistry, 1995. 34(39): p. 12543-52.
22. Kleineke, J., C. Duls, and H.D. Soling, *Subcellular compartmentation of guanine nucleotides and functional relationships between the adenine and guanine nucleotide systems in isolated hepatocytes.* FEBS Lett, 1979. 107(1): p. 198-202.
23. Markby, D.W., R. Onrust, and H.R. Bourne, *Separate GTP binding and GTPase activating domains of a G alpha subunit.* Science, 1993. 262(5141): p. 1895-901.
24. Vetter, I.R. and A. Wittinghofer, *The guanine nucleotide-binding switch in three dimensions.* Science, 2001. 294(5545): p. 1299-304.
25. Zhang, F.L. and P.J. Casey, *Protein prenylation: molecular mechanisms and functional consequences.* Annu Rev Biochem, 1996. 65: p. 241-69.
26. Orita, S., K. Kaibuchi, S. Kuroda, K. Shimizu, H. Nakanishi, and Y. Takai, *Comparison of kinetic properties between two mammalian ras p21 GDP/GTP exchange proteins, ras guanine nucleotide-releasing factor and smg GDP dissociation stimulation.* J Biol Chem, 1993. 268(34): p. 25542-6.

27. Wells, C., X. Jiang, S. Gutowski, and P.C. Sternweis, *Functional characterization of p115 RhoGEF.* Methods Enzymol, 2002. 345: p. 371-82.
28. Glomset, J.A. and C.C. Farnsworth, *Role of protein modification reactions in programming interactions between ras-related GTPases and cell membranes.* Annu Rev Cell Biol, 1994. 10: p. 181-205.
29. Hori, Y., A. Kikuchi, M. Isomura, M. Katayama, Y. Miura, H. Fujioka, K. Kaibuchi, and Y. Takai, *Post-translational modifications of the C-terminal region of the rho protein are important for its interaction with membranes and the stimulatory and inhibitory GDP/GTP exchange proteins.* Oncogene, 1991. 6(4): p. 515-22.
30. Bhattacharyya, R. and P.B. Wedegaertner, *Galpha 13 requires palmitoylation for plasma membrane localization, Rho-dependent signaling, and promotion of p115-RhoGEF membrane binding.* J Biol Chem, 2000. 275(20): p. 14992-9.
31. Thompson, J.D., D.G. Higgins, and T.J. Gibson, *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Res, 1994. 22(22): p. 4673-80.
32. Bishop, A.L. and A. Hall, *Rho GTPases and their effector proteins.* Biochem J, 2000. 348 Pt 2: p. 241-55.
33. Joneson, T. and D. Bar-Sagi, *A Rac1 effector site controlling mitogenesis through superoxide production.* J Biol Chem, 1998. 273(29): p. 17991-4.
34. Wu, W.J., R. Lin, R.A. Cerione, and D. Manor, *Transformation activity of Cdc42 requires a region unique to Rho-related proteins.* J Biol Chem, 1998. 273(27): p. 16655-8.
35. Wu, W.J., D.A. Leonard, A.C. R, and D. Manor, *Interaction between Cdc42Hs and RhoGDI is mediated through the Rho insert region.* J Biol Chem, 1997. 272(42): p. 26153-8.

36. Kreck, M.L., J.L. Freeman, A. Abo, and J.D. Lambeth, *Membrane association of Rac is required for high activity of the respiratory burst oxidase.* Biochemistry, 1996. 35(49): p. 15683-92.
37. Lebowitz, P.F., J.P. Davide, and G.C. Prendergast, *Evidence that farnesyltransferase inhibitors suppress Ras transformation by interfering with Rho activity.* Mol Cell Biol, 1995. 15(12): p. 6613-22.
38. Hall, A., *Rho GTPases and the actin cytoskeleton.* Science, 1998. 279(5350): p. 509-14.
39. Sah, V.P., T.M. Seasholtz, S.A. Sagi, and J.H. Brown, *The role of Rho in G protein-coupled receptor signal transduction.* Annu Rev Pharmacol Toxicol, 2000. 40: p. 459-89.
40. Cheng, L., G.M. Mahon, E.V. Kostenko, and I.P. Whitehead, *PH domain-mediated activation of the Rho-specific guanine nucleotide exchange factor Dbs by Rac1.* J Biol Chem, 2003.
41. Bar-Sagi, D. and A. Hall, *Ras and Rho GTPases: a family reunion.* Cell, 2000. 103(2): p. 227-38.
42. Sahai, E. and C.J. Marshall, *RHO-GTPases and cancer.* Nat Rev Cancer, 2002. 2(2): p. 133-42.
43. Eva, A. and S.A. Aaronson, *Isolation of a new human oncogene from a diffuse B-cell lymphoma.* Nature, 1985. 316(6025): p. 273-5.
44. Venter, J.C., et al., *The sequence of the human genome.* Science, 2001. 291(5507): p. 1304-51.
45. Schmidt, A. and A. Hall, *Guanine nucleotide exchange factors for Rho GTPases: turning on the switch.* Genes Dev, 2002. 16(13): p. 1587-609.
46. Hardt, W.D., L.M. Chen, K.E. Schuebel, X.R. Bustelo, and J.E. Galan, *S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells.* Cell, 1998. 93(5): p. 815-26.
47. De Vries, L., B. Zheng, T. Fischer, E. Elenko, and M.G. Farquhar, *The regulator of G protein signaling family.* Annu Rev Pharmacol Toxicol, 2000. 40: p. 235-71.

48. Taya, S., N. Inagaki, H. Sengiku, H. Makino, A. Iwamatsu, I. Urakawa, K. Nagao, S. Kataoka, and K. Kaibuchi, *Direct interaction of insulin-like growth factor-1 receptor with leukemia-associated RhoGEF.* J Cell Biol, 2001. 155(5): p. 809-20.
49. Hart, M.J., S. Sharma, N. elMasry, R.G. Qiu, P. McCabe, P. Polakis, and G. Bollag, *Identification of a novel guanine nucleotide exchange factor for the Rho GTPase.* J Biol Chem, 1996. 271(41): p. 25452-8.
50. Rumenapp, U., A. Blomquist, G. Schworer, H. Schablowski, A. Psoma, and K.H. Jakobs, *Rho-specific binding and guanine nucleotide exchange catalysis by KIAA0380, a dbl family member.* FEBS Lett, 1999. 459(3): p. 313-8.
51. Swiercz, J.M., R. Kuner, J. Behrens, and S. Offermanns, *Plexin-B1 directly interacts with PDZ-RhoGEF/LARG to regulate RhoA and growth cone morphology.* Neuron, 2002. 35(1): p. 51-63.
52. Kourlas, P.J., M.P. Strout, B. Becknell, M.L. Veronese, C.M. Croce, K.S. Theil, R. Krahe, T. Ruutu, S. Knuutila, C.D. Bloomfield, and M.A. Caligiuri, *Identification of a gene at 11q23 encoding a guanine nucleotide exchange factor: evidence for its fusion with MLL in acute myeloid leukemia.* Proc Natl Acad Sci U S A, 2000. 97(5): p. 2145-50.
53. Rowley, J.D., *The critical role of chromosome translocations in human leukemias.* Annu Rev Genet, 1998. 32: p. 495-519.
54. Reuther, G.W., Q.T. Lambert, M.A. Booden, K. Wennerberg, B. Becknell, G. Marcucci, J. Sondek, M.A. Caligiuri, and C.J. Der, *Leukemia-associated Rho guanine nucleotide exchange factor, a Dbl family protein found mutated in leukemia, causes transformation by activation of RhoA.* J Biol Chem, 2001. 276(29): p. 27145-51.
55. Suzuki, N., S. Nakamura, H. Mano, and T. Kozasa, *Galpha 12 activates Rho GTPase through tyrosine-phosphorylated leukemia-associated RhoGEF.* Proc Natl Acad Sci U S A, 2003. 100(2): p. 733-8.

56. Hart, M.J., X. Jiang, T. Kozasa, W. Roscoe, W.D. Singer, A.G. Gilman, P.C. Sternweis, and G. Bollag, *Direct stimulation of the guanine nucleotide exchange activity of p115 RhoGEF by Galpha13.* Science, 1998. 280(5372): p. 2112-4.

57. Wells, C.D., S. Gutowski, G. Bollag, and P.C. Sternweis, *Identification of potential mechanisms for regulation of p115 RhoGEF through analysis of endogenous and mutant forms of the exchange factor.* J Biol Chem, 2001. 276(31): p. 28897-905.

58. Wells, C.D., M.Y. Liu, M. Jackson, S. Gutowski, P.M. Sternweis, J.D. Rothstein, T. Kozasa, and P.C. Sternweis, *Mechanisms for reversible regulation between G13 and Rho exchange factors.* J Biol Chem, 2002. 277(2): p. 1174-81.

59. Strathmann, M.P. and M.I. Simon, *G alpha 12 and G alpha 13 subunits define a fourth class of G protein alpha subunits.* Proc Natl Acad Sci U S A, 1991. 88(13): p. 5582-6.

60. Boquet, P. and E. Lemichez, *Bacterial virulence factors targeting Rho GTPases: parasitism or symbiosis?* Trends Cell Biol, 2003. 13(5): p. 238-46.

61. Buhl, A.M., N.L. Johnson, N. Dhanasekaran, and G.L. Johnson, *G alpha 12 and G alpha 13 stimulate Rho-dependent stress fiber formation and focal adhesion assembly.* J Biol Chem, 1995. 270(42): p. 24631-4.

62. Kozasa, T., X. Jiang, M.J. Hart, P.M. Sternweis, W.D. Singer, A.G. Gilman, G. Bollag, and P.C. Sternweis, *p115 RhoGEF, a GTPase activating protein for Galpha12 and Galpha13.* Science, 1998. 280(5372): p. 2109-11.

63. Chikumi, H., S. Fukuhara, and J.S. Gutkind, *Regulation of G protein-linked guanine nucleotide exchange factors for Rho, PDZ-RhoGEF, and LARG by tyrosine phosphorylation: evidence of a role for focal adhesion kinase.* J Biol Chem, 2002. 277(14): p. 12463-73.

64. Booden, M.A., D.P. Siderovski, and C.J. Der, *Leukemia-associated Rho guanine nucleotide exchange factor promotes G alpha q-coupled activation of RhoA.* Mol Cell Biol, 2002. 22(12): p. 4053-61.

65. Sunahara, R.K., J.J. Tesmer, A.G. Gilman, and S.R. Sprang, *Crystal structure of the adenylyl cyclase activator Gsalpha.* Science, 1997. 278(5345): p. 1943-7.
66. Coleman, D.E., A.M. Berghuis, E. Lee, M.E. Linder, A.G. Gilman, and S.R. Sprang, *Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis.* Science, 1994. 265(5177): p. 1405-12.
67. Lambright, D.G., J. Sondek, A. Bohm, N.P. Skiba, H.E. Hamm, and P.B. Sigler, *The 2.0 A crystal structure of a heterotrimeric G protein.* Nature, 1996. 379(6563): p. 311-9.
68. Kozasa, T., *Purification of Recombinant G Protein α and $\beta\gamma$ Subunits from Sf9 Cells*, in *G Proteins. Techniques of Analysis*, D. Manning, Editor. 1999, CRC Press LLC: Boca Raton, Florida. p. 23-38.
69. Bach, H., Y. Mazor, S. Shaky, A. Shoham-Lev, Y. Berdichevsky, D.L. Gutnick, and I. Benhar, *Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies.* J Mol Biol, 2001. 312(1): p. 79-93.
70. Watson, N., M.E. Linder, K.M. Druey, J.H. Kehrl, and K.J. Blumer, *RGS family members: GTPase-activating proteins for heterotrimeric G-protein alpha-subunits.* Nature, 1996. 383(6596): p. 172-5.
71. Georgiou, G. and P. Valax, *Expression of correctly folded proteins in Escherichia coli.* Curr Opin Biotechnol, 1996. 7(2): p. 190-7.
72. Houry, W.A., D. Frishman, C. Eckerskorn, F. Lottspeich, and F.U. Hartl, *Identification of in vivo substrates of the chaperonin GroEL.* Nature, 1999. 402(6758): p. 147-54.
73. Tesmer, J.J., D.M. Berman, A.G. Gilman, and S.R. Sprang, *Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis.* Cell, 1997. 89(2): p. 251-61.
74. Whitehead, I.P., S. Campbell, K.L. Rossman, and C.J. Der, *Dbl family proteins.* Biochim Biophys Acta, 1997. 1332(1): p. F1-23.

75. Pasteris, N.G., A. Cadle, L.J. Logie, M.E. Porteous, C.E. Schwartz, R.E. Stevenson, T.W. Glover, R.S. Wilroy, and J.L. Gorski, *Isolation and characterization of the faciogenital dysplasia (Aarskog-Scott syndrome) gene: a putative Rho/Rac guanine nucleotide exchange factor.* Cell, 1994. 79(4): p. 669-78.

76. Heisterkamp, N., J.R. Stephenson, J. Groffen, P.F. Hansen, A. de Klein, C.R. Bartram, and G. Grosveld, *Localization of the c-ab1 oncogene adjacent to a translocation break point in chronic myelocytic leukaemia.* Nature, 1983. 306(5940): p. 239-42.

77. Zhu, K., B. Debreceni, F. Bi, and Y. Zheng, *Oligomerization of DH domain is essential for Dbl-induced transformation.* Mol Cell Biol, 2001. 21(2): p. 425-37.

78. Bottomley, M.J., K. Salim, and G. Panayotou, *Phospholipid-binding protein domains.* Biochim Biophys Acta, 1998. 1436(1-2): p. 165-83.

79. Blomberg, N., E. Baraldi, M. Nilges, and M. Saraste, *The PH superfold: a structural scaffold for multiple functions.* Trends Biochem Sci, 1999. 24(11): p. 441-5.

80. Yao, L., P. Janmey, L.G. Frigeri, W. Han, J. Fujita, Y. Kawakami, J.R. Apgar, and T. Kawakami, *Pleckstrin homology domains interact with filamentous actin.* J Biol Chem, 1999. 274(28): p. 19752-61.

81. Baumeister, M.A., L. Martinu, K.L. Rossman, J. Sondek, M.A. Lemmon, and M.M. Chou, *Loss of phosphatidylinositol 3-phosphate binding by the C-terminal Tiam-1 pleckstrin homology domain prevents in vivo Rac1 activation without affecting membrane targeting.* J Biol Chem, 2003. 278(13): p. 11457-64.

82. Cheng, L., G.M. Mahon, E.V. Kostenko, and I.P. Whitehead, *Pleckstrin homology domain-mediated activation of the rho-specific guanine nucleotide exchange factor Dbs by Rac1.* J Biol Chem, 2004. 279(13): p. 12786-93.

83. Zheng, Y., D. Zangrilli, R.A. Cerione, and A. Eva, *The pleckstrin homology domain mediates transformation by oncogenic dbl through specific intracellular targeting.* J Biol Chem, 1996. 271(32): p. 19017-20.

84. Lodowski, D.T., J.A. Pitcher, W.D. Capel, R.J. Lefkowitz, and J.J. Tesmer, *Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma.* Science, 2003. 300(5623): p. 1256-62.

85. Bi, F., B. Debreceni, K. Zhu, B. Salani, A. Eva, and Y. Zheng, *Autoinhibition mechanism of proto-Dbl.* Mol Cell Biol, 2001. 21(5): p. 1463-74.

86. Nimnual, A.S., B.A. Yatsula, and D. Bar-Sagi, *Coupling of Ras and Rac guanosine triphosphatases through the Ras exchanger Sos.* Science, 1998. 279(5350): p. 560-3.

87. Han, J., K. Luby-Phelps, B. Das, X. Shu, Y. Xia, R.D. Mosteller, U.M. Krishna, J.R. Falck, M.A. White, and D. Broek, *Role of substrates and products of PI 3-kinase in regulating activation of Rac-related guanosine triphosphatases by Vav.* Science, 1998. 279(5350): p. 558-60.

88. Rossman, K.L., L. Cheng, G.M. Mahon, R.J. Rojas, J.T. Snyder, I.P. Whitehead, and J. Sondek, *Multifunctional roles for the PH domain of Dbs in regulating Rho GTPase activation.* J Biol Chem, 2003. 278(20): p. 18393-400.

89. Crespo, P., K.E. Schuebel, A.A. Ostrom, J.S. Gutkind, and X.R. Bustelo, *Phosphotyrosine-dependent activation of Rac-1 GDP/GTP exchange by the vav proto-oncogene product.* Nature, 1997. 385(6612): p. 169-72.

90. Wang, D.S., T. Deng, and G. Shaw, *Membrane binding and enzymatic activation of a Dbl homology domain require the neighboring pleckstrin homology domain.* Biochem Biophys Res Commun, 1997. 234(1): p. 183-9.

91. Fukuhara, S., H. Chikumi, and J.S. Gutkind, *Leukemia-associated Rho guanine nucleotide exchange factor (LARG) links heterotrimeric G proteins of the G(12) family to Rho.* FEBS Lett, 2000. 485(2-3): p. 183-8.

92. Kristelly, R., B.T. Earnest, L. Krishnamoorthy, and J.J. Tesmer, *Preliminary structure analysis of the DH/PH domains of leukemia-associated RhoGEF.* Acta Crystallogr D Biol Crystallogr, 2003. 59(Pt 10): p. 1859-62.

93. Kapust, R.B., J. Tozser, J.D. Fox, D.E. Anderson, S. Cherry, T.D. Copeland, and D.S. Waugh, *Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency.* Protein Eng, 2001. 14(12): p. 993-1000.
94. Otwinowski, Z. and V. Minor, Methods in Enzymology, 1997. 276: p. 307-326.
95. Vagin, A. and A. Teplyakov, *An approach to multi-copy search in molecular replacement.* Acta Crystallogr D Biol Crystallogr, 2000. 56 Pt 12: p. 1622-4.
96. Kissinger, C.R., D.K. Gehlhaar, and D.B. Fogel, *Rapid automated molecular replacement by evolutionary search.* Acta Crystallogr D Biol Crystallogr, 1999. 55 (Pt 2): p. 484-91.
97. Schwede, T., J. Kopp, N. Guex, and M.C. Peitsch, *SWISS-MODEL: An automated protein homology-modeling server.* Nucleic Acids Res, 2003. 31(13): p. 3381-5.
98. Kleywegt, G.J. and T.A. Jones, *Software for handling macromolecular envelopes.* Acta Crystallogr D Biol Crystallogr, 1999. 55 (Pt 4): p. 941-4.
99. Jones, T.A., J.Y. Zou, S.W. Cowan, and Kjeldgaard, *Improved methods for building protein models in electron density maps and the location of errors in these models.* Acta Crystallogr A, 1991. 47 (Pt 2): p. 110-9.
100. Hendrickson, W.A., J.R. Horton, and D.M. LeMaster, *Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure.* Embo J, 1990. 9(5): p. 1665-72.
101. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography.* Acta Crystallographica D, 1994. 50: p. 760-763.
102. Otwinowski, Z. in *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*. 1991. Warrington: Daresbury Laboratory.

103. Brunger, A.T., P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren, *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr D Biol Crystallogr, 1998. 54 (Pt 5): p. 905-21.
104. Winn, M.D., M.N. Isupov, and G.N. Murshudov, *Use of TLS parameters to model anisotropic displacements in macromolecular refinement.* Acta Crystallogr D Biol Crystallogr, 2001. 57(Pt 1): p. 122-33.
105. Self, A.J. and A. Hall, *Purification of recombinant Rho/Rac/G25K from Escherichia coli.* Methods Enzymol, 1995. 256: p. 3-10.
106. Ihara, K., S. Muraguchi, M. Kato, T. Shimizu, M. Shirakawa, S. Kuroda, K. Kaibuchi, and T. Hakoshima, *Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue.* J Biol Chem, 1998. 273(16): p. 9656-66.
107. Zhu, K., B. Debreceni, R. Li, and Y. Zheng, *Identification of Rho GTPase-dependent sites in the Dbl homology domain of oncogenic Dbl that are required for transformation.* J Biol Chem, 2000. 275(34): p. 25993-6001.
108. Blomberg, N. and M. Nilges, *Functional diversity of PH domains: an exhaustive modelling study.* Fold Des, 1997. 2(6): p. 343-55.
109. Blomberg, N., E. Baraldi, M. Sattler, M. Saraste, and M. Nilges, *Structure of a PH domain from the C. elegans muscle protein UNC-89 suggests a novel function.* Structure Fold Des, 2000. 8(10): p. 1079-87.
110. Isakoff, S.J., T. Cardozo, J. Andreev, Z. Li, K.M. Ferguson, R. Abagyan, M.A. Lemmon, A. Aronheim, and E.Y. Skolnik, *Identification and analysis of PH domain-containing targets of phosphatidylinositol 3-kinase using a novel in vivo assay in yeast.* Embo J, 1998. 17(18): p. 5374-87.

111. Snyder, J.T., K.L. Rossman, M.A. Baumeister, W.M. Pruitt, D.P. Siderovski, C.J. Der, M.A. Lemmon, and J. Sondek, *Quantitative analysis of the effect of phosphoinositide interactions on the function of Dbl family proteins.* J Biol Chem, 2001. 276(49): p. 45868-75.

112. Guex, N. and M.C. Peitsch, *SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling.* Electrophoresis, 1997. 18(15): p. 2714-23.

113. Pruitt, W.M., A.E. Karnoub, A.C. Rakauskas, M. Guipponi, S.E. Antonarakis, A. Kurakin, B.K. Kay, J. Sondek, D.P. Siderovski, and C.J. Der, *Role of the pleckstrin homology domain in intersectin-L Dbl homology domain activation of Cdc42 and signaling.* Biochim Biophys Acta, 2003. 1640(1): p. 61-8.

114. Mistou, M.Y., E. Jacquet, P. Poullet, H. Rensland, P. Gideon, I. Schlichting, A. Wittinghofer, and A. Parmeggiani, *Mutations of Ha-ras p21 that define important regions for the molecular mechanism of the SDC25 C-domain, a guanine nucleotide dissociation stimulator.* Embo J, 1992. 11(7): p. 2391-7.

115. Valencia, A., P. Chardin, A. Wittinghofer, and C. Sander, *The ras protein family: evolutionary tree and role of conserved amino acids.* Biochemistry, 1991. 30(19): p. 4637-48.

116. John, J., R. Sohmen, J. Feuerstein, R. Linke, A. Wittinghofer, and R.S. Goody, *Kinetics of interaction of nucleotides with nucleotide-free H-ras p21.* Biochemistry, 1990. 29(25): p. 6058-65.

117. Rensland, H., J. John, R. Linke, I. Simon, I. Schlichting, A. Wittinghofer, and R.S. Goody, *Substrate and product structural requirements for binding of nucleotides to H-ras p21: the mechanism of discrimination between guanosine and adenosine nucleotides.* Biochemistry, 1995. 34(2): p. 593-9.

118. Debant, A., C. Serra-Pages, K. Seipel, S. O'Brien, M. Tang, S.H. Park, and M. Streuli, *The multidomain protein Trio binds the LAR transmembrane tyrosine phosphatase, contains a protein kinase domain, and has separate rac-specific and rho-specific guanine nucleotide exchange factor domains.* Proc Natl Acad Sci U S A, 1996. 93(11): p. 5466-71.

119. Blangy, A., E. Vignal, S. Schmidt, A. Debant, C. Gauthier-Rouviere, and P. Fort, *TrioGEF1 controls Rac- and Cdc42-dependent cell structures through the direct activation of rhoG.* J Cell Sci, 2000. 113 (Pt 4): p. 729-39.

120. Rojas, R.J., R.J. Kimple, K.L. Rossman, D.P. Siderovski, and J. Sondek, *Established and emerging fluorescence-based assays for G-protein function: Ras-superfamily GTPases.* Comb Chem High Throughput Screen, 2003. 6(4): p. 409-18.

121. Cheng, L., K.L. Rossman, G.M. Mahon, D.K. Worthylake, M. Korus, J. Sondek, and I.P. Whitehead, *RhoGEF specificity mutants implicate RhoA as a target for Dbs transforming activity.* Mol Cell Biol, 2002. 22(19): p. 6895-905.

122. Bellanger, J.M., S. Estrach, S. Schmidt, A. Briancon-Marjollet, O. Zugasti, S. Fromont, and A. Debant, *Different regulation of the Trio Dbl-Homology domains by their associated PH domains.* Biol Cell, 2003. 95(9): p. 625-34.

123. Tan, Y.C., H. Wu, W.N. Wang, Y. Zheng, and Z.X. Wang, *Characterization of the interactions between the small GTPase RhoA and its guanine nucleotide exchange factors.* Anal Biochem, 2002. 310(2): p. 156-62.

124. Schlumberger, M.C., A. Friebel, G. Buchwald, K. Scheffzek, A. Wittinghofer, and W.D. Hardt, *Amino acids of the bacterial toxin SopE involved in G nucleotide exchange on Cdc42.* J Biol Chem, 2003. 278(29): p. 27149-59.

125. Compton, S.J. and C.G. Jones, *Mechanism of dye response and interference in the Bradford protein assay.* Anal Biochem, 1985. 151(2): p. 369-74.

126. Wiechelman, K.J., R.D. Braun, and J.D. Fitzpatrick, *Investigation of the bicinchoninic acid protein assay: identification of the groups responsible for color formation.* Anal Biochem, 1988. 175(1): p. 231-7.

127. Zheng, Y., J.A. Glaven, W.J. Wu, and R.A. Cerione, *Phosphatidylinositol 4,5-bisphosphate provides an alternative to guanine nucleotide exchange factors by stimulating the dissociation of GDP from Cdc42Hs*. J Biol Chem, 1996. 271(39): p. 23815-9.

VITA

Romana Kristelly was born in Vienna, Austria, on September 24, 1972, the daughter of Gertraud and Werner Kristelly. After graduating from Auf der Schmelz High School, Vienna, Austria in 1990, she entered the College of Pharmacy at the University of Vienna, Austria. She received her *Magistra of Pharmacy* degree in February 1997. From May 1997 to May 1998 Romana was working as a post-graduate fellow of the Austrian Chamber of Community Pharmacists in the Pharmacy *Zur Maria Heil der Kranken* in Vienna, Austria. She passed her community pharmacy registration exam in May 1998 and worked as a community pharmacist until October 1998. In January 1999 she was admitted to the Graduate School of The University of Texas at Austin. Romana was married to Marcus on July 25th 1999 and their daughter Anika was born on June 11th 2003.

Permanent Address: Trazerberggasse 65B/2, 1130 Wien, Austria (Europe).

This dissertation was typed by the author.